



```

PR 07-JUN-1995; 95US-0487748.
PR 03-MAR-1995; 95US-0398633.
XX (MILL-) MILLENNIUM PHARM INC.
XX PA
XX PI
XX Levinson DA;
XX WPI: 1996-433404/43.
XX N-PSDB; AAT38265.
XX
XX Genes and their products differentially expressed in T helper cells
XX - useful in diagnosis and treatment of immune disorders, e.g.
XX multiple sclerosis, asthma, lepromatous leprosy, etc.
XX
XX Claim 10; Fig 17; 218pp; English.
XX
XX The product (AAW01047) of novel murine gene 200 (AAT38265) is a novel
XX receptor contg. an extracellular Ig domain. Gene 200 expression is
XX many-fold higher in TH1 than in TH2 cell subpopulations.
XX
XX Modulation of the 200 gene product may ameliorate a range of
XX T-cell-related disorders. Soluble gene 200 products (e.g. fusions
XX to immunoglobulins) can be produced that increase the blood
XX half-life of the product. Transgenic animals expressing the 200
XX gene product are useful models of TH cell subpopulation-related
XX disorders. The human homologue (AAW01049) of the gene 200 product
XX has also been identified.
XX
XX Sequence 281 AA;
SQ
Query Match 100.0%; Score 1468; DB 17; Length 281;
Best Local Similarity 100.0%; Pred. No. 2,1e-124;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGGLTNCVLLLLQLLARSLEDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFPCW 60
DB 1 MSGGLTNCVLLLLQLLARSLEDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFPCW 60
QY 61 SOCTNELRTDERNTYKSSRYOLKGDINKGVSLIKNTVLDHGYCCRIQFGLMN 120
DB 61 SOCTNELRTDERNTYKSSRYOLKGDINKGVSLIKNTVLDHGYCCRIQFGLMN 120
QY 121 DKLELKLDIRAKAYTPAQTAGDSTTASPTLTTERNGSETOVLVTLHNNNGTKISTWA 180
DB 121 DKLELKLDIRAKAYTPAQTAGDSTTASPTLTTERNGSETOVLVTLHNNNGTKISTWA 180
QY 181 DEIKSGETIRTAIHIGVGSAGLTLALIIIGVLIKMYSCKKKSLSLITLANLPFG 240
DB 181 DEIKSGETIRTAIHIGVGSAGLTLALIIIGVLIKMYSCKKKSLSLITLANLPFG 240
QY 241 LANAGAVRIRSENIYTIENVEVENSNEYCYNNSQOPS 281
DB 241 LANAGAVRIRSENIYTIENVEVENSNEYCYNNSQOPS 281
RESULT 2
AAY97056
ID AAY97056 standard; Protein; 281 AA.
XX
XX AAY97056;
XX
XX 31-OCT-2000 (first entry)
XX
XX Murine T helper cell differentially expressed gene 200 product.
XX
XX T helper cell; differential expression; 200 gene; immunomodulator;
XX anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;
XX chymotryptic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;
XX protozoacide; lymphocyte; modulator; gene therapy.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX Peptide 1..20

```

```

FT Protein 21..281
FT /label= Mature_protein
FT Domain 21..192
FT /label= Extracellular_domain
FT Domain 193..214
FT /label= Transmembrane_domain
FT Domain 215..208
FT /label= Cytoplasmic_domain
XX
XX US6084083-A.
XX
XX 04-JUL-2000.
XX
XX 28-MAR-1997; 97US-0829525.
XX
XX 01-MAR-1996; 96US-0609583.
XX 03-MAR-1995; 95US-0398633.
XX 07-JUN-1995; 95US-0487748.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Levinson DA;
XX
XX WPI: 2000-464385/40.
XX N-PSDB; AAM51898.
XX
XX New isolated human 200 gene products or polypeptides, useful for
XX treating and diagnosing immune disorders, especially T helper
XX lymphocyte-related disorders
XX
XX Example; Fig 17A-D; 107pp; English.
XX
XX Genes which are differentially expressed within and among T helper (TH)
XX cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,
XX can be used diagnostically or as targets for therapeutic intervention.
XX The polypeptides are useful for treating and diagnosing of immune
XX disorders, especially T lymphocyte-related disorders. These disorders
XX include chronic inflammatory diseases and disorders (e.g. Crohn's
XX disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or
XX Grave's disease), or atopic conditions (e.g. asthma and allergy,
XX including allergic rhinitis or food allergies). Also included are
XX certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.
XX HIV) or bacterial (e.g. tuberculosis or lepromatous leprosy) infections.
XX
XX Sequence 281 AA;
SQ
Query Match 100.0%; Score 1468; DB 21; Length 281;
Best Local Similarity 100.0%; Pred. No. 2,1e-124;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGGLTNCVLLLLQLLARSLEDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFPCW 60
DB 1 MSGGLTNCVLLLLQLLARSLEDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFPCW 60
QY 61 SOCTNELRTDERNTYKSSRYOLKGDINKGVSLIKNTVLDHGYCCRIQFGLMN 120
DB 61 SOCTNELRTDERNTYKSSRYOLKGDINKGVSLIKNTVLDHGYCCRIQFGLMN 120
QY 121 DKLELKLDIRAKAYTPAQTAGDSTTASPTLTTERNGSETOVLVTLHNNNGTKISTWA 180
DB 121 DKLELKLDIRAKAYTPAQTAGDSTTASPTLTTERNGSETOVLVTLHNNNGTKISTWA 180
QY 181 DEIKSGETIRTAIHIGVGSAGLTLALIIIGVLIKMYSCKKKSLSLITLANLPFG 240
DB 181 DEIKSGETIRTAIHIGVGSAGLTLALIIIGVLIKMYSCKKKSLSLITLANLPFG 240
QY 241 LANAGAVRIRSENIYTIENVEVENSNEYCYNNSQOPS 281
DB 241 LANAGAVRIRSENIYTIENVEVENSNEYCYNNSQOPS 281
RESULT 3
AAM50225

```



ID AAM50225 standard; Protein: 281 AA.  
 XX  
 AC AAM50225;  
 XX  
 DT 07-JUN-2002 (first entry)  
 XX  
 DE Mouse 200 gene product, preferentially expressed in TH1 cells.  
 XX  
 KM Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;  
 KM receptor; differential expression; immune disorder; psoriasis;  
 KM multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
 KM antiproliferative; diagnosis; therapy.  
 XX  
 OS Mus musculus.  
 XX  
 PN US6288218-B1.  
 XX  
 PD 11-SEP-2001.  
 XX  
 PF 25-SEP-1997; 97US-0937339.  
 XX  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (LEVI/) LEVINSON D A.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI; 2001-647189/74.  
 DR N-PSDB; AAI70263, AAI70255.  
 XX  
 PT Detecting 200 gene expression in a sample, useful for treating and  
 PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
 PT comprises detecting the presence of a 200 gene product or an RNA  
 PT encoding the 200 gene product -  
 XX  
 PS Claim 2(a); Fig 17A-B; 108pp; English.  
 XX  
 CC The present sequence is that of the protein product of the mouse  
 CC 200 gene (see AAI70263). The 200 gene is preferentially expressed  
 CC in mature, fully differentiated T helper subpopulation TH1 cells  
 CC relative to cells. The gene can be used diagnostically or as  
 CC a target for therapeutic intervention for the treatment of immune  
 CC disorders. A claimed method for detecting 200 gene expression in  
 CC a sample involves detecting the level of a 200 gene product, or an  
 CC RNA encoding it. The gene product detected may be the present  
 CC amino acid sequence, or amino acids 1-20, 1-192, 1-214, 21-192,  
 CC 21-214, 21-281, 293-214, 193-281 or 215-281 of it. Detection of  
 CC the human 200 gene product (see AAM50223) provides a means of  
 CC diagnosing a TH1 cell subpopulation related immune disorder, such  
 CC as multiple sclerosis, psoriasis or insulin-dependent diabetes  
 CC (claimed). In addition to the 200 gene, the invention provides  
 CC other genes that are differentially expressed within and among  
 CC TH cells and TH cell subpopulations and which can be used in  
 CC methods for the diagnosis, prognosis, evaluation and treatment of  
 CC TH cell subpopulation-related disorders, for the identification of  
 CC subjects exhibiting a predisposition to such conditions, for  
 CC monitoring patients undergoing clinical evaluation for the  
 CC treatment of such disorders, and for monitoring the efficacy of  
 CC compounds used in clinical trials. Other immune disorders that can  
 CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
 CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
 CC Grave's disease, contact dermatitis, graft rejection, graft versus  
 CC host disease, sarcoidosis, atopic conditions, asthma, allergy,  
 CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
 CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
 CC viral infection (e.g. HIV), and bacterial infection (e.g.  
 CC tuberculosis and lepromatous leprosy).  
 XX  
 SQ Sequence 281 AA;  
 Query Match 100.0%; Score 1468; DB 22; Length 281;

Best Local Similarity 100.0%; Pred. No. 2,1e-124;  
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSGLTNCVLLTLLQILLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60  
 DB 1 MFSGLTNCVLLTLLQILLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCKGKCPW 60  
 QY 61 SGCNELLRTDERENVYQKSSRYQLKGDNLKGDVSLIKVNTLDDHGTCCRIOPGLAM 120  
 DB 61 SGCNELLRTDERENVYQKSSRYQLKGDNLKGDVSLIKVNTLDDHGTCCRIOPGLAM 120  
 QY 121 DKKLELKLIDIKAAKVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
 DB 121 DKKLELKLIDIKAAKVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
 QY 181 DEIKDGETTIRTAHIGVGSAGITLALLIGVILKMSCKKKLSLSITLNLPPGG 240  
 DB 181 DEIKDGETTIRTAHIGVGSAGITLALLIGVILKMSCKKKLSLSITLNLPPGG 240  
 QY 241 LANAGAVRIRSEENITYTEENVYEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENITYTEENVYEVENSNEYCYVNSQOPS 281  
 RESULT 4  
 AAB81511  
 ID AAB81511 standard; Protein: 281 AA.  
 AC AAB81511;  
 XX  
 DT 18-JUN-2001 (first entry)  
 XX  
 DE Murine TH1 specific 200 gene product.  
 XX  
 KM Murine; T helper cell; TH cell; TH1; TH2; immunomodulator;  
 KM anti-inflammatory; antiallergic; dermatological; antiviral;  
 KM antibacterial; T helper lymphocyte modulator; gene therapy;  
 KM TH specific gene; 200 gene; immune disorder; inflammation;  
 XX  
 OS Mus sp.  
 XX  
 PN US6204371-B1.  
 XX  
 PD 20-MAR-2001.  
 XX  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI; 2001-272703/28.  
 DR N-PSDB; AAF82609.  
 XX  
 PT New murine or human 200 genes and their corresponding polypeptides,  
 PT useful for treating or diagnosing immune disorders, especially T helper  
 PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies  
 PT -  
 XX  
 PS Claim 5; Fig 17; 109pp; English.  
 XX  
 CC The present sequence is encoded by murine 200 gene, which is expressed  
 CC at higher levels in T helper (TH1) cells than in TH2 cells. The  
 CC invention relates to an isolated nucleic acid molecule that comprises  
 CC the full length murine 200 gene or full length human 200 gene nucleotide  
 CC sequence. The nucleic acids are useful for treating or diagnosing immune  
 CC disorders, especially T helper lymphocyte-related disorders, e.g.  
 CC inflammatory diseases (e.g. Crohn's disease), multiple sclerosis, Grave's  
 CC disease, contact dermatitis, psoriasis, asthma and allergies, or certain

CC viral (e.g. HIV) or bacterial (e.g. tuberculosis) infections.  
XX  
SQ Sequence 281 AA;  
Query Match 100.0%; Score 1468; DB 22; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMSGTTLNCVLLDQLLLARSLIEDGYKVEGKNAVLPCSYTLPTSGTLVPMCGKGFPCW 60  
DB 1 MMSGTTLNCVLLDQLLLARSLIEDGYKVEGKNAVLPCSYTLPTSGTLVPMCGKGFPCW 60  
QY 61 SCSCTHELRTDERNVTYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 SCSCTHELRTDERNVTYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
QY 121 DKLELKLDIRAAYTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKLDIRAAYTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLSLITLANLPFGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLSLITLANLPFGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
RESULT 5  
AAB50696  
ID AAB50696 strand; Protein; 281 AA.  
AC AAB50696;  
XX  
XX 20-MAR-2001 (first entry)  
DE Mouse 200 gene product amino acid sequence SEQ ID NO:10.  
XX  
XX Treatment; diagnosis; immune disorder; mast cell related disorder;  
KW T-helper lymphocyte-related disorder; ischaemic disorder;  
KW identification; vasodilator; cardiac; antianginal; angina pectoris;  
KW ischaemic renal disease; myocardial ischaemia; myocardial infarction;  
KW cortical infarction; ischaemic injury; kidney transplant.  
XX  
OS Mus musculus.  
XX  
XX WO200073498-A1.  
XX  
XX 07-DEC-2000.  
XX  
XX 31-MAY-2000; 2000WO-US14986.  
XX  
XX 02-JUN-1999; 99US-0324986.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Levinson DA, Lloyd CM, McCarthy SA;  
PI  
XX  
XX WPI; 2001-016510/02.  
DR N-PSDB; AAC90986.  
XX  
XX Ameliorating a symptom of an ischemic disorder or injury in a mammal  
PT e.g. ischemic renal disease or myocardial ischemia, by administering a  
PT 200 gene product (S1), a nucleic acid encoding (S1) or an antibody  
PT directed against (S1) -  
XX  
XX Claim 10; Fig 17; 309pp; English.  
XX  
XX The present invention describes a method for ameliorating a symptom of  
CC an ischaemic disorder or injury in a mammal. The method comprises  
CC administering a 200 gene product, a nucleic acid encoding (S1) or an  
CC antibody directed against (S1). The method is useful for treating a

CC symptom of an ischaemic disorder such as ischaemic renal disease or  
CC myocardial ischaemia (such as angina pectoris), myocardial or cortical  
CC infarction. The method is also useful for treating a symptom of an  
CC ischaemic injury occurring due to transplantation of a kidney. The  
CC present sequence is used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 281 AA;  
Query Match 100.0%; Score 1468; DB 22; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMSGTTLNCVLLDQLLLARSLIEDGYKVEGKNAVLPCSYTLPTSGTLVPMCGKGFPCW 60  
DB 1 MMSGTTLNCVLLDQLLLARSLIEDGYKVEGKNAVLPCSYTLPTSGTLVPMCGKGFPCW 60  
QY 61 SCSCTHELRTDERNVTYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
DB 61 SCSCTHELRTDERNVTYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIOPGLMN 120  
QY 121 DKLELKLDIRAAYTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKLDIRAAYTPAQTAAHGDSTTASPTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLSLITLANLPFGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTALIIIGVILIKWYSCKKKLSLSLITLANLPFGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
RESULT 6  
AAB59162  
ID AAB59162 standard; protein; 281 AA.  
AC AAB59162;  
XX  
XX 21-MAR-2001 (first entry)  
DE Murine 200 gene protein.  
XX  
XX Cysteine protease; immune disorder; T lymphocyte; Crohn's;  
KW arthritis; diabetes; multiple sclerosis; viral infection; bacterial;  
KW HIV.  
XX  
XX Mus sp..  
XX  
XX US6156887-A.  
XX  
XX 05-DEC-2000.  
XX  
XX 03-OCT-1997; 97US-0939729.  
XX  
XX 01-MAR-1996; 96US-0609583.  
XX  
XX 03-MAR-1995; 95US-0398633.  
XX  
XX 07-JUN-1995; 95US-0487748.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Levinson DA;  
PI  
XX  
XX WPI; 2001-101473/11.  
DR  
XX  
XX Novel polypeptide exhibiting cysteine protease activity, useful for  
PT treating and diagnosing immune disorders, especially T  
PT lymphocyte-related disorders, e.g. Crohn's disease, multiple sclerosis,  
PT graft versus host disease or allergies -  
XX  
XX Examples; Fig 17; 107pp; English.  
XX

Query Match	Best Local Similarity	100.0%	Score 1468	DB 22	Length 281	Matches 281	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MFSGTLNLCVLLLLDQLLLARSLEDEGKYKEVGKNAVLPSCYTLPTSGTLVPMCMGKGFPM	60							
Db	1	MFSGTLNLCVLLLLDQLLLARSLEDEGKYKEVGKNAVLPSCYTLPTSGTLVPMCMGKGFPM	60							
Qy	61	SOCTNELLRTERNAVITYKSSRYQKGLNKGDSVLIKNVTLDDHGTCCRIQPEGLMN	120							
Db	61	SOCTNELLRTERNAVITYKSSRYQKGLNKGDSVLIKNVTLDDHGTCCRIQPEGLMN	120							
Qy	121	DKLEELKIDIAKAKTPAQTAAAGDSSTASPRLLTERNGSEFQTLVTLANNNGTKISTWA	180							
Db	121	DKLEELKIDIAKAKTPAQTAAAGDSSTASPRLLTERNGSEFQTLVTLANNNGTKISTWA	180							
Qy	181	DEIKDSGETIRTAIHIGVVSAGLTALAIIGVLIKMYSCCKKKKSSSLITLANLPDGG	240							
Db	181	DEIKDSGETIRTAIHIGVVSAGLTALAIIGVLIKMYSCCKKKKSSSLITLANLPDGG	240							
Qy	241	LANAGAVIRSEENITYTTEENYEVENSNEYCYNQQPS	281							
Db	241	LANAGAVIRSEENITYTTEENYEVENSNEYCYNQQPS	281							
RESULT 7										
ABG32767	ABG32767	standard; Protein; 281 AA.								
XX	XX	ABG32767;								
AC	AC	15-NOV-2002 (first entry)								
DT	DT	Mouse TH1 specific 200 protein.								
XX	XX	TH1; TH2; antiinflammatory; antidiabetic; antithyroid; atopy;								
KW	KW	antiaesthetic; anti-allergic; virucide; antiarthritic; dermatological;								
KW	KW	antipneumatic; nephrotoxic; immunosuppressive; immune response;								
KW	KW	asthma; allergy; allergic rhinitis; viral infection; thyroiditis;								
KW	KW	inflammatory disease; Crohn's disease; arthritis; diabetes; mouse;								
KW	KW	dermatitis; psoriasis; glomerular nephritis; autoimmunity;								
KW	KW	graft rejection.								
OS	OS	Mus sp.								
XX	XX	US6414117-B1.								
PN	PN	02-JUL-2002.								
PD	PD	12-MAY-1999;	99US-0310367.							
XX	XX	01-MAR-1996;	96US-0609583.							
XX	XX	28-MAR-1997;	97US-0828525.							
PR	PR	03-MAR-1995;	95US-0398633.							
PR	PR	07-JUN-1995;	95US-0487748.							
XX	XX	(MILL-) MILLENNIUM PHARM INC.								

XX	Levinson DA;
F1	WP1; 2002-641576/69.
DR	
XX	
PT	Novel polypeptides which are differentially expressed within and among
PT	T helper cells and cell populations, useful for treating immune
PT	disorders, especially T helper cell subpopulation-related disorders -
XX	
PS	Example; Fig 17, 109pp; English.
XX	
CC	This invention relates to a novel isolated polypeptide which is a T
CC	helper (Th1) cell subpopulation specific gene product. The invention
CC	also discloses other Th1 or Th2 specific gene products. The proteins of
CC	the invention may have antiinflammatory, antidiabetic, antithyroid,
CC	antiasmatic, antiallergic, virocidic, antiarthritic, dermatological,
CC	antiapoptotic, nephrotropic and immunosuppressive activities and may be
CC	used as a regulator of the immune response. The proteins of the
CC	invention are also useful to reduce the level of Th2 cell activity for
CC	treating Th1 cell subpopulation-related disorders including atopic
CC	conditions, such as asthma and allergy including allergic rhinitis, the
CC	effects of pathogen, including viral infection, chronic inflammatory
CC	diseases such as Crohn's disease, arthritis, diabetes, thyroiditis,
CC	dermatitis, psoriasis, glomerular nephritis, organ-specific
CC	autoimmunity, graft rejection and graft versus host disease. The
CC	present sequence represents the gene 200 protein isolated from
CC	a Th1/Th2 differential display library, this protein is specifically
CC	expressed in Th1 cells.
XX	
SQ	Sequence 281 AA;
	Query Match 100.0%; Score 1469; DB 23; Length 281;
	Best Local Similarity 100.0%; Pred. No. 2.le-124;
	Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MFSGITNCVLLIQLLARSLEFGYKVEGKNAYLPSCSYTLPRISGLIVPMCKKGPCPW 60
DB	1 MFSGITNCVLLIQLLARSLEFGYKVEGKNAYLPSCSYTLPRISGLIVPMCKKGPCPW 60
OY	61 SGCNTELLRTDERNAVTVOKSSRYOLAKGDINKGVSLIIKNVTLPDHGTGCCRIOPGLMN 120
DB	61 SGCNTELLRTDERNAVTVOKSSRYOLAKGDINKGVSLIIKNVTLPDHGTGCCRIOPGLMN 120
OY	121 DKCLELKDIKAAYTPAQTAGHSDTSPPTLTTERNGSETOTLVTLNNNGTKISTWA 180
DB	121 DKCLELKDIKAAYTPAQTAGHSDTSPPTLTTERNGSETOTLVTLNNNGTKISTWA 180
OY	181 DEIDSGETTITAHIIGVGVSAGITALLIIGVILLIKMYSCKKKLSLSLTTLANLPDGG 240
DB	181 DEIDSGETTITAHIIGVGVSAGITALLIIGVILLIKMYSCKKKLSLSLTTLANLPDGG 240
OY	241 LANGAVALRRESENIYTIEENVYEVENSNEYCYVNSQQPS 281
DB	241 LANGAVALRRESENIYTIEENVYEVENSNEYCYVNSQQPS 281
RESULT 8	
ABP70434	
ID	ABP70434 standard; Protein; 281 AA.
AC	ABP70434;
XX	
DT	22-APR-2003 (first entry)
XX	
DE	Amino acid sequence of murine TIM-3 BAIB/c allele.
XX	
KW	T cell immunoglobulin domain, mucin domain, TIM-1, TIM-2, TIM-3;
KW	TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW	myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW	allergic T cell response; autoimmune disease.
OS	Mus musculus.
XX	

PN W02003002722-A2.  
XX  
XX 09-JAN-2003.  
XX  
XX 01-JUL-2002; 2002WO-US20890.  
XX  
XX 29-JUN-2001; 2001US-302344P.  
XX  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
PI Mcintire JF, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
XX WPI: 2003-210268/20.  
DR N-PSDB; AB668328.  
XX  
XX New nucleic acid comprising a mammalian T cell Immunoglobulin domain  
PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
PT allergy, eczema or autoimmune disease -  
XX  
XX Claim 10; Page 75; 94pp; English.  
XX  
XX The present sequence is a murine T cell immunoglobulin domain and  
CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
CC conserved IGV and mucin domains. The locus comprising the TIM family is  
CC genetically associated with immune dysfunction, including asthma. The  
CC TIM gene family is located within a region of human chromosome 5 that  
CC is commonly deleted in malignancies and myelodysplastic syndrome.  
CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
CC airway hyperactivity and allergic T cell responses, and other variants  
CC associated with protection against these responses. T cells express TIM  
CC proteins, which critically regulate CD4 T cell differentiation. Th1  
CC cells preferentially express TIM-3, while Th2 cells preferentially  
CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
XX  
XX Sequence 281 AA:  
SQ  
Query Match 100.0%; Score 1468; DB 24; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTNCVLLLOLLARSLLEDGKVEGKNAVPCSYTPTSGLVPMCGKGFPM 60  
DB 1 MFSGLTNCVLLLOLLARSLLEDGKVEGKNAVPCSYTPTSGLVPMCGKGFPM 60  
QY 61 SGTNEELLRTDERNVYQKSSRYQKGDVSLIKNVTLDHGYCCRIQFPGLMN 120  
DB 61 SGTNEELLRTDERNVYQKSSRYQKGDVSLIKNVTLDHGYCCRIQFPGLMN 120  
QY 121 DKLEKLKIDIKAAKTPAQTAGDSTTASPRLLTTERNSSETQTLVLTNNNGTISTWA 180  
DB 121 DKLEKLKIDIKAAKTPAQTAGDSTTASPRLLTTERNSSETQTLVLTNNNGTISTWA 180  
QY 181 DEIKDSGETIRTAHIGVSVAGLTLALITGLILKVCCKKKLSLTLTANLPQGG 240  
DB 181 DEIKDSGETIRTAHIGVSVAGLTLALITGLILKVCCKKKLSLTLTANLPQGG 240  
QY 241 IANAGAVRIRSEENYITIEENYEVENSNEYCYVNSQOPS 281  
DB 241 IANAGAVRIRSEENYITIEENYEVENSNEYCYVNSQOPS 281  
XX  
XX RESULT 9  
XX ABG73716 standard; Protein; 281 AA.  
XX ID ABG73716;  
XX AC ABG73716;  
XX XX  
XX 09-APR-2003 (first entry)  
XX XX  
XX Murine TH1-associated 200 protein.  
XX DE  
XX XX

KW T-cell receptor; TH; T helper cell; 103 gene; TH2 cell; murine;  
KW TH2 cell marker; TH1 cell; protozoacide; antibacterial; virucide;  
KW immunosuppressive; antiinflammatory; antiarthritic; antidiabetic;  
KW neuroprotective; dermatological; antihypertensive; antiparasitic; helminthic;  
KW nephrotoxic; antiaesthetic antiallergic; CD8 agonist; CD4 agonist;  
KW interleukin agonist; bacterial; viral infection; immune disorder;  
KW Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;  
KW Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;  
KW graft rejection; graft versus host disease; asthma; glomerulonephritis;  
KW allergy; gene therapy; TH cell subpopulation.  
XX  
XX Mus musculus.  
XX  
XX US6455685-B1.  
XX  
XX 24-SEP-2002.  
XX  
XX 27-FEB-1998; 98US-0032337.  
XX  
XX 03-MAR-1995; 95US-0398633.  
XX 07-JUN-1995; 95US-0487748.  
XX 01-MAR-1996; 96US-0609583.  
XX  
XX (MILL-) MILENNIUM PHARM INC.  
XX  
XX Levinson DA;  
XX  
XX WPI: 2003-066247/06.  
XX N-PSDB; ABQ77037; ABQ77038.  
XX  
XX Identifying a test compound that binds to a 103 gene product or is a  
PT TH2 cell marker, useful for the identification and therapeutic use of  
PT compounds as treatments of helminthic, bacterial and viral infections,  
PT and immune disorders -  
XX  
XX Example 8; Figure 17A-D; 131pp; English.  
XX  
XX This invention describes a novel method for identifying a test compound  
CC that binds to a 103 gene product or is a TH2 cell marker. The method  
CC comprises contacting a test compound with an immobilised 103 gene  
CC product, removing unbound test compound or separating the complex from  
CC the reaction mixture, and detecting the complex. Identifying a test  
CC compound that binds to a 103 gene product alternatively comprises: (a)  
CC contacting a test compound with a cell engineered to express a 103 gene  
CC product or co-expressing a 103 gene product and a test compound in a  
CC cell; and (b) removing unbound 103 gene product. Identifying a test  
CC compound that is a TH2 cell marker further comprises detecting a complex,  
CC where contacting the test compound with a TH2 and TH1 cell for the test  
CC compound to bind either cell and detecting binding where it indicates the  
CC test compound as a TH2 cell marker. The products described in the  
CC invention have protozoacide, antibacterial, virucide, immunosuppressive,  
CC antiinflammatory, antiaesthetic, antidiabetic, neuroprotective,  
CC dermatological, antihypertensive, antiparasitic, antiaesthetic;  
CC and antiallergic activity and can act as CD8, CD4 and interleukin  
CC agonists. The methods and compositions of the present invention are  
CC useful for the identification and therapeutic use of compounds as  
CC treatments of helminthic, bacterial and viral infections and immune  
CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
CC asthma, allergy and glomerulonephritis. They can also be used for gene  
CC therapy, for the diagnostic evaluation and prognosis of TH cell  
CC subpopulation-related disorders, identification of subjects exhibiting a  
CC predisposition to such conditions, monitoring undergoing clinical  
CC evaluation and efficacy for the treatment of the disorders. This sequence  
CC represents a T-helper cell associated polypeptide described in the  
CC disclosure of the invention.  
XX  
XX Sequence 281 AA:  
SQ  
Query Match 100.0%; Score 1468; DB 24; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2.1e-124;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



PT multiple sclerosis, asthma, lepromatous leprosy, etc.  
XX  
PS Claim 10; Fig 24; 218bp; English.  
XX  
CC The product (AA01049) of novel human gene 200 (AA738267) is a novel  
CC cell surface receptor of the Ig superfamily class. Gene 200  
CC expression is many-fold higher in TH1 than in TH2 subpopulations.  
CC Modulation of the 200 gene product may ameliorate a range of  
CC T-cell-related disorders. Soluble gene 200 products (e.g. fusions  
CC to immunoglobulins) can be produced that increase the blood  
CC half-life of the product. Transgenic animals expressing the 200  
CC gene product are useful models of TH cell subpopulation-related  
CC disorders. The murine homologue (AA01047) of the gene 200 product  
CC has also been identified.  
XX  
SQ Sequence 301 AA;  
XX  
Query Match 60.1%; Score 883; DB 17; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
QY 1 MFGSLTNCVLLQLLARSLEDGYKVEGNAYLPCSYTLPTSGTLVPMCGKGFCEW 60  
DB 1 MFSLPFDVCLLLLLLRSSSEVEYRAVGONALPCFYTPAAPGNLVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDVSLIKNTLDDHGTCCRIOPRGLMN 120  
DB 61 FEGGNVLTBTDERDNYW-TSRYWLNGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLDIKAAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDSGETIRTAIHIGVVSAGLTALIIIGVLIKWSCKKKKLSLSL 230  
DB 180 ANELDRSLANDLRSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONSL 239  
QY 231 ITLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNS-QOOS 281  
DB 240 ISLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNSROOPS 291  
RESULT 12  
AA97058  
ID AA97058 standard; Protein; 301 AA.  
XX  
AC AA97058;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE Human T helper cell differentially expressed gene 200 product.  
XX  
XX T helper cell; differential expression; 200 gene; immunomodulator;  
KW anti-inflammatory; anti-arthritis; antibacterial; immunosuppressive;  
KW thymomimetic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;  
KW protozoacide; lymphocyte; modulator; gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1..20 /label= Signal\_peptide  
FT 21..301 /label= Mature\_protein  
FT 21..200 /label= Mature\_protein  
FT 201..224 /label= Extracellular\_domain  
FT 225..301 /label= Transmembrane\_domain  
FT Domain  
XX  
XX US6084083-A.  
XX  
XX

PD 04-JUL-2000.  
XX  
XX 28-MAR-1997; 97US-0829525.  
XX  
XX 01-MAR-1996; 96US-0609583.  
PR 03-MAR-1995; 95US-0398633.  
PR 07-JUN-1995; 95US-0487748.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Levinson DA;  
XX  
DR WPI: 2000-464385/40.  
DR N-PSDB; AAA51904, AAA51905.  
XX  
PT New isolated human 200 gene products or polypeptides, useful for  
PT treating and diagnosing immune disorders, especially T helper  
PT lymphocyte-related disorders  
XX  
PS Claim 1; Fig 24A-D; 107bp; English.  
XX  
XX Genes which are differentially expressed within and among T helper (TH)  
XX cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,  
XX can be used diagnostically or as targets for therapeutic intervention.  
XX The polypeptides are useful for treating and diagnosing of immune  
XX disorders, especially T lymphocyte-related disorders. These disorders  
XX include chronic inflammatory diseases and disorders (e.g. Crohn's  
XX disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or  
XX Grave's disease), or atopic conditions (e.g. asthma and allergy,  
XX including allergic rhinitis or food allergies). Also included are  
XX certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g.  
XX HIV) or bacterial (e.g. tuberculosis or lepromatous leprosy) infections.  
XX  
SQ Sequence 301 AA;  
XX  
Query Match 60.1%; Score 883; DB 21; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
QY 1 MFGSLTNCVLLQLLARSLEDGYKVEGNAYLPCSYTLPTSGTLVPMCGKGFCEW 60  
DB 1 MFSLPFDVCLLLLLLRSSSEVEYRAVGONALPCFYTPAAPGNLVPVCMGKACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDVSLIKNTLDDHGTCCRIOPRGLMN 120  
DB 61 FEGGNVLTBTDERDNYW-TSRYWLNGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLDIKAAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLVIRPAKYTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDSGETIRTAIHIGVVSAGLTALIIIGVLIKWSCKKKKLSLSL 230  
DB 180 ANELDRSLANDLRSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONSL 239  
QY 231 ITLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNS-QOOS 281  
DB 240 ISLANLPBGGLANAGAVIRSEENITYTEENVYEVENSNEYCYVNSROOPS 291  
RESULT 13  
AAM50223  
ID AAM50223 standard; Protein; 301 AA.  
XX  
AC AAM50223;  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Human 200 gene product, preferentially expressed in TH1 cells.  
XX  
XX Human; 200 gene; T helper; T lymphocyte; T cell; TH1;  
KW receptor; differential expression; immune disorder; psoriasis;  
KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
XX  
XX

KW antipsoriatic; diagnosis; therapy.  
XX  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT Protein /label= Signal\_peptide  
FT Protein /label= Mature\_protein  
FT Domain 21..301  
FT Domain /label= Extracellular\_domain  
FT Domain 201..224  
FT Domain /label= Transmembrane\_domain  
FT Domain 225..301  
FT Domain /label= Cytoplasmic\_domain  
XX  
XX US6288218-B1.  
XX  
XX 11-SEP-2001.  
XX  
XX 25-SEP-1997; 97US-0937399.  
XX  
XX 01-MAR-1996; 96US-0609583.  
XX 03-MAR-1995; 95US-0398633.  
XX 07-JUN-1995; 95US-0487748.  
XX  
XX (LEVI/) LEVINSON D A.  
XX  
XX Levinson DA;  
XX  
XX MPI; 2001-647189/74.  
XX N-PSDB; AAI70254, AAI70281.  
XX  
XX Detecting 200 gene expression in a sample, useful for treating and  
PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
PT comprises detecting the presence of a 200 gene product or an RNA  
PT encoding the 200 gene product -  
XX  
XX  
XX Claim 1(a); Fig 24A-B; 108pp; English.  
XX  
XX The present sequence is that of the protein product of the human  
XX 200 gene (see AAI70254). It is a receptor of the Ig superfamily  
XX class. The 200 gene is preferentially expressed in mature, fully  
XX differentiated T helper subpopulation TH1 cells relative to  
XX subpopulation TH2 cells. The gene can be used diagnostically or as  
XX a target for therapeutic intervention for the treatment of immune  
XX disorders. A claimed method for diagnosing a TH cell subpopulation  
XX related immune disorder involves detecting the level of a 200 gene  
XX product, or an RNA encoding it, so that if the level differs from  
XX that in a control sample, the disorder is diagnosed. The gene  
XX product detected may be the present amino acid sequence, or  
XX amino acids 1-20, 1-200, 1-224, 30-128, 21-200, 21-224, 21-301,  
XX 201-224, 201-301 or 224-301 of it. The immune disorder is  
XX especially a TH1 cell subpopulation-related immune disorder, such  
XX as multiple sclerosis, psoriasis or insulin-dependent diabetes  
XX (claimed). In addition to the 200 gene, the invention provides  
XX other genes that are differentially expressed within and among  
XX TH cells and TH cell subpopulations and which can be used in  
XX methods for the diagnosis, prognosis, evaluation and treatment of  
XX TH cell subpopulation-related disorders, for the identification of  
XX subjects exhibiting a predisposition to such conditions, for  
XX monitoring patients undergoing clinical evaluation for the  
XX treatment of such disorders, and for monitoring the efficacy of  
XX compounds used in clinical trials. Other immune disorders that can  
XX be treated/diagnosed include Crohn's disease, reactive arthritis,  
XX Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
XX Grave's disease, contact dermatitis, graft rejection, allergic  
XX host disease, sarcoidosis, atopic conditions, asthma, allergy,  
XX allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
XX glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
XX viral infection (e.g. HIV), and bacterial infection (e.g.  
XX tuberculosis and lepromatous leprosy).

SEQ Sequence 301 AA;  
Query Match 60.1%; Score 883; DB 22; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;  
QY 1 MFSGLTNCVLLLLQILLARSLEDGYKVEGKNAYLPCSYTLPTSGLVPMCKGFCPW 60  
DB 1 MFSHLPFDCVLLLLLLLRSSSEVEYRAVEGQNAVLPFCFTPAAPNLVAVCGKAGCFV 60  
QY 61 SQCTNELLTDERBNTVYQKSSRYQLKGDJNKGDVSLIKNVTLLDHYGTCRQIFGLAN 120  
DB 61 FECGNVLRTRDERDVYVW-TSRWYLNGBDFRKGDSLTLENTVLTADSGIYCCRIQIPIGN 119  
QY 121 DKLEELKLIKAKVTPAQTAFHSDSTTASPRLTTERNG-SENQTLVTLHNNNGTRKSTW 179  
DB 120 DEKFNLKLVTKPKVTPAPLQDFPAAPFRMLTTGHGPAETQJGSLPDINLTQISTL 179  
QY 180 ADE-----IKDSGETIRTAIHIGVVSAGLTALIGVLLIKWYSCKKKLSLSL 230  
DB 180 ANELRDSRLANDRDSGATIRIGIYAGICGALALIFGALIFKWSKSKKIQRLSL 239  
QY 231 ITLANLPFGGLANAGVRIKSEENIYTTIBNRYEVSNSERYCYVNS-QQPS 281  
DB 240 ISLANLPFGGLANAAVEGIRSEENIYTTIBNRYEVEEPNEYCYVSSRQOPS 291  
RESULT 14  
AAU14409  
ID AAU14409 standard; Protein; 301 AA.  
XX  
XX AAU14409;  
AC  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Human novel protein #280.  
XX  
XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cytoprotective; neuroprotective; vulnerary; nocropic;  
KW anticonvulsant; antirheumatic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; haemostatic; antilastmatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200155437-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02623.  
XX  
XX 25-JAN-2001; 2000US-0491404.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Dymnac RT;  
XX  
XX WPI; 2001-451939/48.  
XX N-PSDB; AAS22714.  
XX  
XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage -  
XX  
XX Example 4; Page 804-805; 894pp; English.  
XX  
XX The invention relates to polynucleotides encoding novel human  
XX proteins or their active domains. The polypeptides, polynucleotides and  
XX antibodies raised against the polypeptides are used in a method of  
XX treatment of a mammal and prevention of disorders caused by the aberrant  
XX protein expression or activity. The polypeptides can be used as  
XX molecular weight markers, food supplements, and in antibody production.



CC The polypeptides are used to identify compounds which bind to the  
CC polypeptides. Polynucleotides of the invention are used as probes and  
CC primers, for sequencing, for chromosome or gene mapping, in the  
CC production of recombinant proteins, and in generating anti-sense DNA or  
CC RNA and in gene therapy. Polypeptides of the invention can be used to  
CC target drugs to a tumour, in assays to determine biological activity, to  
CC raise antibodies/ elicit an immune response, to determine quantitative  
CC protein levels, as tissue markers, and to isolate receptors or ligands.  
CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
CC anti-inflammatory diseases, nervous system disorders, and infection.  
CC The present sequence represents a protein of the invention.

SQ Sequence 301 AA;

Query Match 60.1%; Score 883; DB 22; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGITLNCVLLLOLLARSLSDGYKVEGKNAVLPCSYTLPITGTLVPMCKGFCPW 60  
DB 1 MFSLPFCVLLLLLLTLRSSEVYRAEVGNALPCFTYPAAGNLVPCWKGACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDVSLIKNTLDDHGTCCRIOPGILMN 120  
DB 61 FECGVVARTDERDVNW-TSRWMLNGDFRGDVSITENVTLDGSGYCCRIQIPGILMN 119  
QY 121 DKLELKIDIKAAKVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNLIKVIKPAKVTAPPTLQORDFTAAFPRLTLTRGHGPAFTQGLSPINLTQISTL 179  
QY 180 ADE-----IKSGEITRTAHIGVGSAGLTALITGVLLIKWSSCKKKLSL 230  
DB 180 ANELRDSRLANDLRSGATIRIGIYIGAGICAGLALALFGLALFKWYSHSEKIQNLSTL 239  
QY 231 ITLANLPQGLANAGAVRIRSEENIYTIENVEYENSNEYCYVNS--QOPS 281  
DB 240 ISLANLPQGLANAGAVRIRSEENIYTIENVEYENNEYCYVNSRQOPS 291

RESULT 15

AAB93838  
ID AAB93838 standard; Protein; 301 AA.

XX AAB93838;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:13669.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Ota T, Itoai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.

PT primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

PS Claim 8; SEQ ID 13669; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH3166 to AAH13628 and  
CC AAH31633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

SQ Sequence 301 AA;

Query Match 60.1%; Score 883; DB 22; Length 301;  
Best Local Similarity 63.7%; Pred. No. 1.8e-71;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGITLNCVLLLOLLARSLSDGYKVEGKNAVLPCSYTLPITGTLVPMCKGFCPW 60  
DB 1 MFSLPFCVLLLLLLTLRSSEVYRAEVGNALPCFTYPAAGNLVPCWKGACPV 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDVSLIKNTLDDHGTCCRIOPGILMN 120  
DB 61 FECGVVARTDERDVNW-TSRWMLNGDFRGDVSITENVTLDGSGYCCRIQIPGILMN 119  
QY 121 DKLELKIDIKAAKVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNLIKVIKPAKVTAPPTLQORDFTAAFPRLTLTRGHGPAFTQGLSPINLTQISTL 179  
QY 180 ADE-----IKSGEITRTAHIGVGSAGLTALITGVLLIKWSSCKKKLSL 230  
DB 180 ANELRDSRLANDLRSGATIRIGIYIGAGICAGLALALFGLALFKWYSHSEKIQNLSTL 239  
QY 231 ITLANLPQGLANAGAVRIRSEENIYTIENVEYENSNEYCYVNS--QOPS 281  
DB 240 ISLANLPQGLANAGAVRIRSEENIYTIENVEYENNEYCYVNSRQOPS 291

Search completed: November 22, 2003, 05:40:20  
Job time : 58.0412 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:38:42 ; Search time 18.3471 Seconds  
(without alignments)  
648.024 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468

Sequence: 1 MFSGLTINCVLLTLQLLAR.....VEVENSNEYCVNSQPS 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A COMB .pep:.\*  
2: /cgn2\_6/ptodata/2/1aa/5B COMB .pep:.\*  
3: /cgn2\_6/ptodata/2/1aa/6A COMB .pep:.\*  
4: /cgn2\_6/ptodata/2/1aa/6B COMB .pep:.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS COMB .pep:.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pcp:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1468	100.0	281	1 US-08-487-748A-9	Sequence 9, App1
2	1468	100.0	281	1 US-08-487-748A-10	Sequence 10, App1
3	1468	100.0	281	3 US-08-480-070C-10	Sequence 10, App1
4	1468	100.0	281	3 US-08-829-525-10	Sequence 10, App1
5	1468	100.0	281	3 US-08-609-583A-10	Sequence 10, App1
6	1468	100.0	281	3 US-08-937-389-10	Sequence 10, App1
7	1468	100.0	281	4 US-09-310-367-10	Sequence 10, App1
8	1468	100.0	281	4 US-09-032-337-10	Sequence 10, App1
9	1468	100.0	281	4 US-09-464-231-10	Sequence 10, App1
10	883	60.1	301	3 US-08-829-525-24	Sequence 24, App1
11	883	60.1	301	3 US-08-597-495B-24	Sequence 24, App1
12	883	60.1	301	3 US-08-937-389-24	Sequence 24, App1
13	883	60.1	301	4 US-09-310-367-24	Sequence 24, App1
14	883	60.1	301	4 US-09-032-337-24	Sequence 24, App1
15	883	60.1	301	4 US-09-464-231-24	Sequence 24, App1
16	267.5	18.2	451	1 US-08-287-001A-2	Sequence 2, App1
17	267.5	18.2	451	5 PCT-US95-09941-2	Sequence 2, App1
18	157	10.7	319	1 US-08-597-495B-22	Sequence 22, App1
19	157	10.7	319	3 US-09-068-051A-22	Sequence 22, App1
20	157	10.7	319	4 US-09-336-536-67	Sequence 67, App1
21	157	10.7	319	4 US-09-254-465A-6	Sequence 6, App1
22	151	10.3	273	4 US-09-254-465A-26	Sequence 26, App1
23	150	10.2	270	4 US-08-254-465A-24	Sequence 24, App1
24	149.5	10.2	365	4 US-08-928-383B-2	Sequence 2, App1
25	147.5	10.0	352	4 US-09-996-243-505	Sequence 505, App1
26	147.5	10.0	365	2 US-08-979-424-3	Sequence 3, App1
27	147.5	10.0	365	3 US-09-272-496-2	Sequence 2, App1

28	143	9.7	318	3 US-09-068-051A-32	Sequence 32, App1
29	138	9.4	320	3 US-08-205-697A-2	Sequence 2, App1
30	138	9.4	320	3 US-08-702-525-2	Sequence 2, App1
31	138	9.4	320	5 PCT-US95-02576-2	Sequence 2, App1
32	135.5	9.2	466	4 US-09-604-107A-8	Sequence 8, App1
33	131	8.9	398	3 US-09-189-035-6	Sequence 6, App1
34	131	8.9	398	3 US-09-382-086-6	Sequence 6, App1
35	131	8.9	398	4 US-08-999-689A-5	Sequence 5, App1
36	131	8.9	503	4 US-08-938-832-6	Sequence 6, App1
37	128.5	8.8	431	3 US-09-038-832-2	Sequence 2, App1
38	128.5	8.8	431	3 US-09-038-832-4	Sequence 4, App1
39	126.5	8.6	365	3 US-08-928-383B-23	Sequence 23, App1
40	126.5	8.6	365	3 US-08-928-383B-24	Sequence 24, App1
41	126.5	8.6	365	3 US-08-928-383B-26	Sequence 26, App1
42	124	8.4	306	2 US-08-147-772-4	Sequence 4, App1
43	124	8.4	306	2 US-08-456-104-8	Sequence 8, App1
44	124	8.4	306	2 US-08-101-624-25	Sequence 25, App1
45	124	8.4	306	3 US-08-153-262-4	Sequence 4, App1

#### ALIGNMENTS

RESULT 1  
US-08-487-748A-9  
Sequence 9, Application US/08487748A  
Patent No. 5721351  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fenime & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,748A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Cornuzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-487-748A-9  
Query Match 100.0%; Score 1468; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 MFSGLTINCVLLTLQLLARSLDGYKVEYKNAVYPCSTLPTSGTLVPMCKGKCPW 60  
Db 1 MFSGLTINCVLLTLQLLARSLDGYKVEYKNAVYPCSTLPTSGTLVPMCKGKCPW 60  
Cy 61 SGCNELLRTDERNNVTYOKSSRYOLKGLNKGVSLIIKKVTLDDHGTCCRIQPGIANN 120

Db 61 SQTNELLRTERDNRVTVYKSSRYOLKGDINKGDSVLIINVTLLDHDGTYCCRIOPGLMN 120  
QY 121 DKLELKLDIKAAKVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
Db 121 DKLELKLDIKAAKVTPAQTAGDSTTASPRITLTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKSGEITRTAHIGHGVASAGLTALIIIGVLIHKWYSCCKKKLSLTLTLANLPPGG 240  
Db 181 DEIKSGEITRTAHIGHGVASAGLTALIIIGVLIHKWYSCCKKKLSLTLTLANLPPGG 240  
QY 241 LANAGAVRIRSEENITYTEENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENITYTEENVEVENSNEYCYVNSQOPS 281

RESULT 2  
US-08-487-748A-10  
; Sequence 10, Application US/08487748A  
; Patent No. 5721351  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,748A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-487-748A-10

Query Match 100.0%; Score 1468; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 DEIKSGEITRTAHIGHGVASAGLTALIIIGVLIHKWYSCCKKKLSLTLTLANLPPGG 240  
Db 181 DEIKSGEITRTAHIGHGVASAGLTALIIIGVLIHKWYSCCKKKLSLTLTLANLPPGG 240  
QY 241 LANAGAVRIRSEENITYTEENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENITYTEENVEVENSNEYCYVNSQOPS 281

RESULT 3  
US-08-480-070C-10  
; Sequence 10, Application US/08480070C  
; Patent No. 6066498  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,070C  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-024  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-480-070C-10

Query Match 100.0%; Score 1468; DB 3; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 4
US-08-829-525-10
; Sequence 10, Application US/08829525
; Patent No. 6084083
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,525
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-829-525-10

Query Match      100.0%; Score 1468; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSGLTINCVLLLLQLLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCPW 60
DB 1 MFSGLTINCVLLLLQLLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCPW 60
QY 61 SOCNEILRTDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIQFPGILMN 120
DB 61 SOCNEILRTDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIQFPGILMN 120
QY 121 DKKLELKLIDIAKAVTPAQTAGDSTTASPRITLTTERNGSETOTLVTLANNNGTKISTWA 180
DB 121 DKKLELKLIDIAKAVTPAQTAGDSTTASPRITLTTERNGSETOTLVTLANNNGTKISTWA 180
QY 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKMYSCKKKLSLSLITLANLPBG 240
DB 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKMYSCKKKLSLSLITLANLPBG 240
QY 241 LANAGAVRIRSEENIYITEENVYEVENSNEYCYVNSQOPS 281
DB 241 LANAGAVRIRSEENIYITEENVYEVENSNEYCYVNSQOPS 281
```

```
RESULT 5
US-08-609-583A-10
; Sequence 10, Application US/08609583A
; Patent No. 6204371
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,583A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-609-583A-10

Query Match      100.0%; Score 1468; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSGLTINCVLLLLQLLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCPW 60
DB 1 MFSGLTINCVLLLLQLLLARSLEDGYKVEGKNAVLPSCSYTLPTSGTLVPMCMGKGFPCPW 60
QY 61 SOCNEILRTDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIQFPGILMN 120
DB 61 SOCNEILRTDERNNVYQKSSRYOLKGDINKGDVSLIKNVTLDHGTCCRIQFPGILMN 120
QY 121 DKKLELKLIDIAKAVTPAQTAGDSTTASPRITLTTERNGSETOTLVTLANNNGTKISTWA 180
DB 121 DKKLELKLIDIAKAVTPAQTAGDSTTASPRITLTTERNGSETOTLVTLANNNGTKISTWA 180
QY 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKMYSCKKKLSLSLITLANLPBG 240
DB 181 DEIKDSEGTIRTAIHIGVGSAGITLALIIIGVLLIKMYSCKKKLSLSLITLANLPBG 240
QY 241 LANAGAVRIRSEENIYITEENVYEVENSNEYCYVNSQOPS 281
DB 241 LANAGAVRIRSEENIYITEENVYEVENSNEYCYVNSQOPS 281
```

## RESULT 6

US-08-937-399-10  
; Sequence 10, Application US/08937399  
; Patent No. 6288218  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,399  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-937-399-10

Query Match 100.0%; Score 1468; DB 3; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.8e-150;

Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSGTLNLCVLLQLQLARSL	EDGYKVEGKNAVLP	CSYTLPTSGTLVPM	CMGKGF	CPW	60
DB	1	MSGTLNLCVLLQLQLARSL	EDGYKVEGKNAVLP	CSYTLPTSGTLVPM	CMGKGF	CPW	60
QY	61	SOCTNELFTDERNVTYQKSSRYQLKGD	LNGDVSLIKNVTLDDHGYCCRI	OPGLMN			120
DB	61	SOCTNELFTDERNVTYQKSSRYQLKGD	LNGDVSLIKNVTLDDHGYCCRI	OPGLMN			120
QY	121	DKKLEKLDIAAAYTPAQT	AGDSTTASPTLTTERNGSET	OTLVTLHNNNGTKISTWA			180
DB	121	DKKLEKLDIAAAYTPAQT	AGDSTTASPTLTTERNGSET	OTLVTLHNNNGTKISTWA			180
QY	181	DEIKDSGETIRTAHIGV	SAGLTALITGVLLIKWY	CKKKKLSLITLANLP	PGG		240
DB	181	DEIKDSGETIRTAHIGV	SAGLTALITGVLLIKWY	CKKKKLSLITLANLP	PGG		240
QY	241	LANAGAVRIRSEENIY	TIENVEVNSNEYCYVNS	QOPS			281
DB	241	LANAGAVRIRSEENIY	TIENVEVNSNEYCYVNS	QOPS			281

## RESULT 7

US-09-310-367-10  
; Sequence 10, Application US/09310367  
; Patent No. 641417  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/310,367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829,525  
; FILING DATE: 28-MAR-1997  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-081  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-310-367-10

Query Match 100.0%; Score 1468; DB 4; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.8e-150;

Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSGTLNLCVLLQLQLARSL	EDGYKVEGKNAVLP	CSYTLPTSGTLVPM	CMGKGF	CPW	60
DB	1	MSGTLNLCVLLQLQLARSL	EDGYKVEGKNAVLP	CSYTLPTSGTLVPM	CMGKGF	CPW	60
QY	61	SOCTNELFTDERNVTYQKSSRYQLKGD	LNGDVSLIKNVTLDDHGYCCRI	OPGLMN			120
DB	61	SOCTNELFTDERNVTYQKSSRYQLKGD	LNGDVSLIKNVTLDDHGYCCRI	OPGLMN			120
QY	121	DKKLEKLDIAAAYTPAQT	AGDSTTASPTLTTERNGSET	OTLVTLHNNNGTKISTWA			180
DB	121	DKKLEKLDIAAAYTPAQT	AGDSTTASPTLTTERNGSET	OTLVTLHNNNGTKISTWA			180
QY	181	DEIKDSGETIRTAHIGV	SAGLTALITGVLLIKWY	CKKKKLSLITLANLP	PGG		240
DB	181	DEIKDSGETIRTAHIGV	SAGLTALITGVLLIKWY	CKKKKLSLITLANLP	PGG		240
QY	241	LANAGAVRIRSEENIY	TIENVEVNSNEYCYVNS	QOPS			281
DB	241	LANAGAVRIRSEENIY	TIENVEVNSNEYCYVNS	QOPS			281

## RESULT 8

US-09-032-337-10  
; Sequence 10, Application US/09032337  
; Patent No. 6455685  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032.337  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cortuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-016  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-032-337-10

Query Match 100.0%; Score 1468; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTNCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
Db 1 MFSGLTNCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
QY 61 SÖCTNELLRDERNNVYÖKSSRYÖLKGDLNKGDVSLIIKNVTLDDHGTCCRIÖPPLMN 120  
Db 61 SÖCTNELLRDERNNVYÖKSSRYÖLKGDLNKGDVSLIIKNVTLDDHGTCCRIÖPPLMN 120  
QY 121 DKKLELKDIDKAAKAVTPAÖTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
Db 121 DKKLELKDIDKAAKAVTPAÖTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDGETIRTAIHIGVGSAGTLALIIIGVLLIKWYSCKKKLSLITLANLPGG 240  
Db 181 DEIKDGETIRTAIHIGVGSAGTLALIIIGVLLIKWYSCKKKLSLITLANLPGG 240  
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

## RESULT 9

US-09-464-231-10  
; Sequence 10, Application US/09464231  
; Patent No. 6562343  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/464,231  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cortuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-464-231-10

Query Match 100.0%; Score 1468; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFSGLTNCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
Db 1 MFSGLTNCVLLIQLLLARSLEDEGYVEVGKNAVLPSCSYTLPTSGTLVPMCMGKGFCPW 60  
QY 61 SÖCTNELLRDERNNVYÖKSSRYÖLKGDLNKGDVSLIIKNVTLDDHGTCCRIÖPPLMN 120  
Db 61 SÖCTNELLRDERNNVYÖKSSRYÖLKGDLNKGDVSLIIKNVTLDDHGTCCRIÖPPLMN 120  
QY 121 DKKLELKDIDKAAKAVTPAÖTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
Db 121 DKKLELKDIDKAAKAVTPAÖTAHGDSTTASPRTLTTERNGSETÖTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDGETIRTAIHIGVGSAGTLALIIIGVLLIKWYSCKKKLSLITLANLPGG 240  
Db 181 DEIKDGETIRTAIHIGVGSAGTLALIIIGVLLIKWYSCKKKLSLITLANLPGG 240  
QY 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENIYTIENVEVENSNEYCYVNSQOPS 281

## RESULT 10

US-08-829-525-24

; Sequence 24, Application US/08829525

; Patent No. 6084083

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESS: Pennie &amp; Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/829,525

; FILING DATE: 28-MAR-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/609,583

; FILING DATE: 01-MAR-1996

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-081

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 301 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; US-08-829-525-24

Query Match 60.1%; Score 883; DB 3; Length 301;

Best Local Similarity 63.7%; Pred. No. 3.9e-87;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MSGGLTINCVLLLLDQLLLARSLLEDGYKVEVGNKAYLPCSYLTPTSTGLVPMCKGKGFCEW 60  
DB 1 MESHLPFCVCLLLLLLTTRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWKGKACPV 60  
QY 61 SOCTNELRTDERNVYTKSSRYOLKGDINKGVSLLIKNVTLDDHGTCCRIQPPGLNN 120  
DB 61 PECGNVLRITDERDNYM-TSRVWLNQDPRKGDVSLTIENVTLADSGIYCCRIQPPGLNN 119  
QY 121 DKKLELDIKAKVTPAQTAAHGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLKVIKAKVTPAQLQDPTAAPPRLMUTTRGHGPAETQTLGSLPDINLQISTL 179  
QY 180 ADE-----IKSGETITRAIHIGVVSAGLTALIIIGVILKWTSCCKKKLSLSL 230  
DB 180 ANELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSSEKIONLSL 239  
QY 231 ITLANLPGLNAGAVIRISEENITYIEENVYEVNSNEYCYVNS-QQPS 281

Db

240 ISLANLPGLNAGAVIRISEENITYIEENVYEVNSNEYCYVNS-QQPS 291

## RESULT 11

US-08-609-583A-24

; Sequence 24, Application US/08609583A

; Patent No. 6204371

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESS: Pennie &amp; Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/609,583A

; FILING DATE: 01-MAR-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-048

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 301 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; US-08-609-583A-24

Query Match 60.1%; Score 883; DB 3; Length 301;

Best Local Similarity 63.7%; Pred. No. 3.9e-87;

Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MSGGLTINCVLLLLDQLLLARSLLEDGYKVEVGNKAYLPCSYLTPTSTGLVPMCKGKGFCEW 60  
DB 1 MESHLPFCVCLLLLLLTTRSSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWKGKACPV 60  
QY 61 SOCTNELRTDERNVYTKSSRYOLKGDINKGVSLLIKNVTLDDHGTCCRIQPPGLNN 120  
DB 61 PECGNVLRITDERDNYM-TSRVWLNQDPRKGDVSLTIENVTLADSGIYCCRIQPPGLNN 119  
QY 121 DKKLELDIKAKVTPAQTAAHGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNKLKVIKAKVTPAQLQDPTAAPPRLMUTTRGHGPAETQTLGSLPDINLQISTL 179  
QY 180 ADE-----IKSGETITRAIHIGVVSAGLTALIIIGVILKWTSCCKKKLSLSL 230  
DB 180 ANELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSSEKIONLSL 239  
QY 231 ITLANLPGLNAGAVIRISEENITYIEENVYEVNSNEYCYVNS-QQPS 281



Db 240 ISLANIPPSGLANAVAGIRSEENITYIENNYEVEEPNEYCYVSSROOPS 291

RESULT 12  
US-08-937-399-24  
; Sequence 24, Application US/08937399  
; Patent No. 6288218

GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,399  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-937-399-24

Query Match 60.1%; Score 883; DB 3; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSGLTINCVALLIQILLARSLEDGYKVEYKNAVLPSCSYTLPTSGTLVPMCKGKFCPW 60  
Db 1 MFSHLPDVCVLLILLTLTRSSSEVEYRAEVQONAVLPFCFYTPAAPGNLVPCWCKGKACP 60  
Qy 61 SOCNELLRTERDEVNTYOKSSRYQLKGDINKGVSLIKKVTLLDDHGTVCRRIOFPGLMN 120  
Db 61 PEGGNVLRTERDEVNTW-TSRVWLNDFPKGVDVSLTEVTLADSGIYCCRIQIPGIMN 119  
Qy 121 DKLELKLIDIKAAKVTPTAQTAGDSTTASPRITLTERNG-SETOTLVTLHNNNGTKISTW 179  
Db 120 DEKNLKLVIKPAKVTAPTLODFTAFPRMLTTRHGGAPEOTIGSLPDINLTQISTL 179  
Qy 180 ADE-----IDSGETIRTAIHIGVGSAGLTALIIIVLILKKNYSCCKKSLSL 230  
Db 180 ANELRDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALLFKWYSHSEKIONLSTL 239

Qy 231 ITLANIPPSGLANAGAVIRSEENITYIENNYEVENSNERYCYVNS-DOOPS 281  
Db 240 ISLANIPPSGLANAVAGIRSEENITYIENNYEVEEPNEYCYVSSROOPS 291

RESULT 13  
US-09-310-367-24  
; Sequence 24, Application US/09310367  
; Patent No. 6414117

GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/310,367  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-310-367-24

Query Match 60.1%; Score 883; DB 4; Length 301;  
Best Local Similarity 63.7%; Pred. No. 3.9e-87;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSGLTINCVALLIQILLARSLEDGYKVEYKNAVLPSCSYTLPTSGTLVPMCKGKFCPW 60  
Db 1 MFSHLPDVCVLLILLTLTRSSSEVEYRAEVQONAVLPFCFYTPAAPGNLVPCWCKGKACP 60  
Qy 61 SOCNELLRTERDEVNTYOKSSRYQLKGDINKGVSLIKKVTLLDDHGTVCRRIOFPGLMN 120  
Db 61 PEGGNVLRTERDEVNTW-TSRVWLNDFPKGVDVSLTEVTLADSGIYCCRIQIPGIMN 119  
Qy 121 DKLELKLIDIKAAKVTPTAQTAGDSTTASPRITLTERNG-SETOTLVTLHNNNGTKISTW 179  
Db 120 DEKNLKLVIKPAKVTAPTLODFTAFPRMLTTRHGGAPEOTIGSLPDINLTQISTL 179

Qy 180 DE-----IKSGEITRTAHIGVGVASGLTALITGVILKPKYSCKKKUSLSL 230  
 Db 180 AMELFDRSLANDRDSGATIRIGTIGAGICMELATLIFGALIFKMTYSIKKXNLSL 239  
 Qy 231 ITLANLPQGLAMAGAVRIRSEENITYIEENYEVENSNEYCYNS-QOPS 281  
 Db 240 ISLANLPQGLANAAVAGIRSEENITYIEENYEVENPEPKYCYVSSROPS 291

RESULT 14  
IIS-09-032

Sequence 24, Application US/09032337  
Patent No. 6455685  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,337  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-032-337-24

Query	Match	Similarity	60.1%;	Score	883;	DB	4;	Length	301;
Best	Local	Similarity	63.7%;	Pred.	No. 3.9e-87;				
Matches	186;	Conservative	26;	Mismatches	68;	Indels	12;	Gaps	4
Qy	1	MFSGLTLCVLLLLDQLLLARSLEDGYKXVGNKAAVLPCTSLPTSGTLVPMCKGKFCPW	60						
Db	1	MFSHLPDCCVLLLLLLLTTRSEVEYRAVGVGNALPCPYTPAAPGNLVPVCMKGKACV	60						
Qy	61	SGCNELLRTERRANTYQKSSRYOLKGLDNKDVSLIKONTLDDHGYTCGRIQFGLNN	120						
Db	61	EECGVAVLTERDRDNYM-TSRMYLNGDPRKQDVSLTIENVTLAOSGIYCCRIQPIGINN	119						
Qy	121	DKLELKKDIIAAKTPAQTAAAGDSVTASPRLLTTERNG-SETQTLVTLAHNNNGTISTW	179						

[illegible]

RESULT 15  
US-09-464

```

1 Sequence 2: Application US/09464231
2 Patent No. 6562343
3 GENERAL INFORMATION:
4 APPLICANT: Levinson, Douglas A.
5 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
6 TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
7 NUMBER OF SEQUENCES: 37
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Penlie & Edmonds
10 STREET: 1155 Avenue of the Americas
11 CITY: New York
12 STATE: New York
13 COUNTRY: USA
14 ZIP: 10036/2711
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Diskette
17 COMPUTER: IBM Compatible
18 OPERATING SYSTEM: DOS
19 SOFTWARE: FastSEQ Version 2.0
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/09/464,231
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/08/609,583
26 FILING DATE: 01-MAR-1996
27 APPLICATION NUMBER: US 08/487,748
28 FILING DATE: 07-JUN-1995
29 APPLICATION NUMBER: US 08/398,633
30 FILING DATE: 03-MAR-1995
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Coruzzi, Laura A.
33 REGISTRATION NUMBER: 30,742
34 REFERENCE/DOCKET NUMBER: 7853-048
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 212-790-9090
37 TELEFAX: 212-869-8864
38 TELEX: 66141 PENNIE
39 INFORMATION FOR SEQ ID NO: 24:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 301 amino acids
42 TYPE: amino acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45 MOLECULE TYPE: protein
46 FRAGMENT TYPE: internal
47 JS-09-464-231-24

```

[illegible]

```

Qy 121 DKLELKDIDIAKAVTPTAQTAHGDSTASPRITLTERNG-SETQTLVTLHNNNGTKISTW 179
Db 120 DEKFNLKLVTKPAKVTPAPTLQORDFTAPPRMLTTRGHGPAETQTLGSLPDINLTQISTL 179
Qy 180 ADE-----IKDSEITRTAHIGVGSAGITLALLIGVLLKWSCKKKLSLSL 230
Db 180 ANELRDSRLANDLRDSGATTIRIGYIGAGICAGLALIFGALIFKWSHSHKEKIQNTLSL 239
Qy 231 ITLANLPPGGLANAGAVRIRSEENITYIENVEVENSNERYCYVNS-OOPS 281
Db 240 ISLANLPPSGLANAVAGIRSEENITYIENVEVEEPNBYCYVSSROOPS 291

```

Search completed: November 22, 2003, 05:45:19  
 Job time : 25.3471 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:43:31 ; Search time 42.0052 Seconds

(without alignments)  
1221.260 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468  
Sequence: 1 MFSLTLNVCVLLQLLLAR.....VVEVENSNEYCYVNSOOPS 281

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1468	100.0	281	US-10-252-131-10	Sequence 10, Appl
2	1468	100.0	281	US-10-004-633-10	Sequence 10, Appl
3	1468	100.0	281	US-10-188-012-9	Sequence 9, Appl
4	1427	97.2	281	US-10-188-012-11	Sequence 11, Appl
5	883	60.1	301	US-10-252-131-24	Sequence 24, Appl
6	883	60.1	301	US-10-004-633-24	Sequence 24, Appl
7	883	60.1	301	US-10-188-012-29	Sequence 29, Appl
8	883	60.1	301	US-10-188-012-31	Sequence 31, Appl
9	520	35.4	185	US-09-529-063-34	Sequence 34, Appl
10	430.5	29.3	142	US-10-414-378-34	Sequence 34, Appl
11	430.5	29.3	142	US-09-966-546-22	Sequence 22, Appl
12	430.5	29.3	142	US-09-966-546-22	Sequence 22, Appl
13	430.5	29.3	142	US-09-965-212-22	Sequence 22, Appl
14	430.5	29.3	142	US-10-189-940-22	Sequence 22, Appl
15	430.5	29.3	142	US-10-189-940-146	Sequence 146, App

16	374	25.5	282	US-10-188-012-3	Sequence 3, Appl
17	367.5	25.0	125	US-10-189-940-145	Sequence 145, App
18	364.5	24.8	305	US-10-188-012-1	Sequence 1, Appl
19	328.5	22.4	359	US-10-188-012-17	Sequence 17, Appl
20	326.5	22.2	359	US-10-188-012-19	Sequence 19, Appl
21	326	22.2	305	US-10-188-012-5	Sequence 5, Appl
22	326	22.2	364	US-10-188-012-7	Sequence 7, Appl
23	324	22.1	364	US-10-188-012-25	Sequence 25, Appl
24	323.5	22.0	365	US-10-188-012-25	Sequence 21, Appl
25	321.5	21.9	359	US-10-188-012-23	Sequence 23, Appl
26	316	21.5	364	US-10-188-012-27	Sequence 27, Appl
27	276.5	18.8	345	US-10-188-012-13	Sequence 13, Appl
28	275	18.7	345	US-10-188-012-15	Sequence 15, Appl
29	274.5	18.7	183	US-09-739-907-65	Sequence 65, Appl
30	274.5	18.7	378	US-10-188-012-33	Sequence 33, Appl
31	274.5	18.7	378	US-10-188-012-35	Sequence 35, Appl
32	274.5	18.7	379	US-09-813-153-138	Sequence 138, App
33	157	10.7	319	US-09-813-153-138	Sequence 6, Appl
34	157	10.7	319	US-09-553-499-6	Sequence 6, Appl
35	157	10.7	319	US-10-265-542-6	Sequence 6, Appl
36	154.5	10.5	323	US-09-971-798-31	Sequence 31, Appl
37	154.5	10.5	343	US-09-971-798-27	Sequence 27, Appl
38	151	10.3	273	US-09-953-499-26	Sequence 26, Appl
39	151	10.3	273	US-10-265-542-26	Sequence 26, Appl
40	150	10.2	268	US-10-265-542-24	Sequence 24, Appl
41	150	10.2	270	US-09-953-499-24	Sequence 24, Appl
42	148	10.1	505	US-10-114-153-12	Sequence 12, Appl
43	147.5	10.0	352	US-09-989-722-505	Sequence 505, App
44	147.5	10.0	352	US-09-989-723-505	Sequence 505, App
45	147.5	10.0	352	US-09-989-279-505	Sequence 505, App

#### ALIGNMENTS

RESULT 1  
US-10-252-131-10  
Sequence 10, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-252-131-10

Query Match 100.0%; Score 1468; DB 12; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
DB 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

#### RESULT 2

US-10-004-633-10  
Sequence 10, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,563  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-633-10

Query Match 100.0%; Score 1468; DB 15; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60

QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
DB 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

#### RESULT 3

US-10-188-012-9  
Sequence 9, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruyff, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
PRIOR FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3 BALB/c allele  
US-10-188-012-9

Query Match 100.0%; Score 1468; DB 15; Length 281;  
Best Local Similarity 100.0%; Pred. No. 5.2e-127;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
DB 1 MSGSLTNCVLLLLQLLARSLSDGYKVEGKNAYLPCSYTLPTSGLVPMCGKGFCEW 60  
QY 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
DB 61 SOCTNELRTDERNTYOKSSRYOLKGDINKGDSLLIKNVTLDHGYCCRIQPGGLMN 120  
QY 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
DB 121 DKLELKDIDIAKAVTPAQTAGDSTTASPRTLTTERNGSETQTLVTLHNNNGTKISTWA 180  
QY 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
DB 181 DEIKDSGETIRTAIHIGVGSAGLTLLALIGVLIKWSCKKKKLSLITLANLPFGG 240  
QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

#### RESULT 4

US-10-188-012-11

Sequence 11, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruff, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3, C.D2 ES-HBA and DBA/2J allele  
US-10-188-012-11

Query Match 97.2%; Score 1427; DB 15; Length 281;  
Best Local Similarity 97.5%; Pred. No. 3,1e-123;  
Matches 274; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MFSGLTNCVLLIQLLLARSLEDGYKVEGKNAVPCSYTLPTSGTLVPMCKGKGCPCW 60  
Db 1 MFSGLTNCVLLIQLLLARSLEDGYKVEGKNAVPCSYTLPTSGTLVPMCKGKGCPCW 60  
QY 61 SQTNELLRTERDENVYQKSRVYOLKGDNLKGDVSLIKVNTLDDHGTCCRIQPGIMN 120  
Db 61 SQTNELLRTERDENVYQKSRVYOLKGDNLKGDVSLIKVNTLDDHGTCCRIQPGIMN 120  
QY 121 DKLELKLIDIAKAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTW 180  
Db 121 DKLELKLIDIAKAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTW 180  
QY 181 DEIKDSEGTIRTAHIGVGSAGITLALIIIGVLLKMYSCCKKKLSLITLTLANLPBG 240  
Db 181 DEIKDSEGTIRTAHIGVGSAGITLALIIIGVLLKMYSCCKKKLSLITLTLANLPBG 240  
QY 241 LANAGAVRIRSEENIYTIENVEVENSENEYCYVNSQOPS 281  
Db 241 LANAGAVRIRSEENIYTIENVEVENSENEYCYVNSQOPS 281

RESULT 5  
US-10-252-131-24  
Sequence 24, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-10-252-131-24

Query Match 60.1%; Score 883; DB 12; Length 301;  
Best Local Similarity 63.7%; Pred. No. 4,4e-73;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSGLTNCVLLIQLLLARSLEDGYKVEGKNAVPCSYTLPTSGTLVPMCKGKGCPCW 60  
Db 1 MFSGLTNCVLLIQLLLARSLEDGYKVEGKNAVPCSYTLPTSGTLVPMCKGKGCPCW 60  
QY 61 SQTNELLRTERDENVYQKSRVYOLKGDNLKGDVSLIKVNTLDDHGTCCRIQPGIMN 120  
Db 61 SQTNELLRTERDENVYQKSRVYOLKGDNLKGDVSLIKVNTLDDHGTCCRIQPGIMN 120  
QY 121 DKLELKLIDIAKAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTW 179  
Db 121 DKLELKLIDIAKAVTPAQTADHGSTASPTLTTERNGSETOTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKDSEGTIRTAHIGVGSAGITLALIIIGVLLKMYSCCKKKLSL 230  
Db 180 ADE-----IKDSEGTIRTAHIGVGSAGITLALIIIGVLLKMYSCCKKKLSL 230  
QY 231 ITLANLPBGGLANAGAVRIRSEENIYTIENVEVENSENEYCYVNS-QOPS 281  
Db 231 ITLANLPBGGLANAGAVRIRSEENIYTIENVEVENSENEYCYVNS-QOPS 281  
QY 240 ISLANLPBGGLANAGAVRIRSEENIYTIENVEVENSENEYCYVNSRQOPS 291  
Db 240 ISLANLPBGGLANAGAVRIRSEENIYTIENVEVENSENEYCYVNSRQOPS 291

RESULT 6  
US-10-004-633-24  
Sequence 24, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748



;; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 24  
;; LENGTH: 301  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-004-633-24

Query Match 60.1%; Score 883; DB 15; Length 301;  
Best Local Similarity 63.7%; Pred. No. 4.4e-73;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MMSGTLNVCVLLLLLOLLARSLSDGYKVEGKNAYLPCSYTLPTSGTLVPMCKGKGFPM 60  
DB 1 MESHLPFDCVLLLLLLLRSSSEVEYRAVGQNAVLPCTYTPAAPGNLVPCMGKACPV 60  
QY 61 SOCTNELARTDERNVTYKSSRYOLKGDNLKGDVSLIKNVTLDDHGTCCRIOPPGMLN 120  
DB 61 FECGNVAVLRTERDNYW-TSRVWLNQDFRKGDSVLTLENVTLADSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLDIRAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNLKLVIPAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKSGETIRTAIHIGVVSAGLTALIIIGVLIKWTSCCKKLSLSL 230  
DB 180 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSKEKIQNLSTL 239  
QY 231 ITLANLPPGGLANAGAVRIRSEENITYTEENVEEVENSEYCYVNS-QOPS 281  
DB 240 ISLANLPPSGLANAVAGIRSEENITYTEENVEEVEEPNEYCYVSSROOPS 291

RESULT 7  
US-10-188-012-29

;; Sequence 29, Application US/10188012  
;; Publication No. US20030124114A1  
;; GENERAL INFORMATION:  
;; APPLICANT: McInliffe, Jennifer Jones  
;; APPLICANT: Umetau, Dale T.  
;; APPLICANT: Dekruyff, Rosemarie  
;; APPLICANT: Kuchroo, Vijay  
;; APPLICANT: Freeman, Gordon J.  
;; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
;; FILE REFERENCE: STAN-235  
;; CURRENT APPLICATION NUMBER: US/10/188,012  
;; PRIOR APPLICATION NUMBER: 2002-07-01  
;; PRIOR FILING DATE: 2001-06-29  
;; NUMBER OF SEQ ID NOS: 36  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 29  
;; LENGTH: 301  
;; TYPE: PRT  
;; ORGANISM: H. sapiens  
;; FEATURE:  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)...(301)  
;; OTHER INFORMATION: TIM-3, allele 1  
US-10-188-012-29

Query Match 60.1%; Score 883; DB 15; Length 301;  
Best Local Similarity 63.7%; Pred. No. 4.4e-73;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MMSGTLNVCVLLLLLOLLARSLSDGYKVEGKNAYLPCSYTLPTSGTLVPMCKGKGFPM 60  
DB 1 MESHLPFDCVLLLLLLLRSSSEVEYRAVGQNAVLPCTYTPAAPGNLVPCMGKACPV 60

QY 61 SOCTNELARTDERNVTYKSSRYOLKGDNLKGDVSLIKNVTLDDHGTCCRIOPPGMLN 120  
DB 61 FECGNVAVLRTERDNYW-TSRVWLNQDFRKGDSVLTLENVTLADSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLDIRAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNLKLVIPAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKSGETIRTAIHIGVVSAGLTALIIIGVLIKWTSCCKKLSLSL 230  
DB 180 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSKEKIQNLSTL 239  
QY 231 ITLANLPPGGLANAGAVRIRSEENITYTEENVEEVENSEYCYVNS-QOPS 281  
DB 240 ISLANLPPSGLANAVAGIRSEENITYTEENVEEVEEPNEYCYVSSROOPS 291

RESULT 8  
US-10-188-012-31

;; Sequence 31, Application US/10188012  
;; Publication No. US20030124114A1  
;; GENERAL INFORMATION:  
;; APPLICANT: McInliffe, Jennifer Jones  
;; APPLICANT: Umetau, Dale T.  
;; APPLICANT: Dekruyff, Rosemarie  
;; APPLICANT: Kuchroo, Vijay  
;; APPLICANT: Freeman, Gordon J.  
;; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
;; FILE REFERENCE: STAN-235  
;; CURRENT APPLICATION NUMBER: US/10/188,012  
;; PRIOR APPLICATION NUMBER: 2002-07-01  
;; PRIOR FILING DATE: 2001-06-29  
;; NUMBER OF SEQ ID NOS: 36  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 31  
;; LENGTH: 301  
;; TYPE: PRT  
;; ORGANISM: H. sapiens  
;; FEATURE:  
;; NAME/KEY: VARIANT  
;; LOCATION: (1)...(301)  
;; OTHER INFORMATION: TIM-3, allele 2  
US-10-188-012-31

Query Match 60.1%; Score 883; DB 15; Length 301;  
Best Local Similarity 63.7%; Pred. No. 4.4e-73;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MMSGTLNVCVLLLLLOLLARSLSDGYKVEGKNAYLPCSYTLPTSGTLVPMCKGKGFPM 60  
DB 1 MESHLPFDCVLLLLLLLRSSSEVEYRAVGQNAVLPCTYTPAAPGNLVPCMGKACPV 60  
QY 61 SOCTNELARTDERNVTYKSSRYOLKGDNLKGDVSLIKNVTLDDHGTCCRIOPPGMLN 120  
DB 61 FECGNVAVLRTERDNYW-TSRVWLNQDFRKGDSVLTLENVTLADSGIYCCRIQIPGIMN 119  
QY 121 DKLELKLDIRAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
DB 120 DEKFNLKLVIPAKVTPAQTAGDSTTASPRLLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ADE-----IKSGETIRTAIHIGVVSAGLTALIIIGVLIKWTSCCKKLSLSL 230  
DB 180 ANELRDSRLANDLRSGATIRIGIYGAGICAGLALALIFGALIFKWTSHSKEKIQNLSTL 239  
QY 231 ITLANLPPGGLANAGAVRIRSEENITYTEENVEEVENSEYCYVNS-QOPS 281  
DB 240 ISLANLPPSGLANAVAGIRSEENITYTEENVEEVEEPNEYCYVSSROOPS 291

RESULT 9  
US-09-529-063-34

; Sequence 34, Application US/09529063  
; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: 058769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; CURRENT FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-063-34

Query Match 35.4%; Score 520; DB 10; Length 185;  
Best Local Similarity 62.1%; Pred. No. 6,1e-40;  
Matches 110; Conservative 15; Mismatches 50; Indels 2; Gaps 2;

QY 1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCYTLPTSGTLVPMCKGKPCPW 60  
DB 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAVGONAYLPFCFYTPAAGNLVPCWKGKACPV 60  
QY 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120  
DB 61 PECGNVLRTERDENVYV-TSRVWLNDFPKGDVSLITENVTLADSGIYCCRIQIPGLMN 119  
QY 121 DKLELKLIDIKAAVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKI 176  
DB 120 DEKNFLKLVKPAKVPAPLTQRDFTAFPRMLTTRGHGPAETQTLGSLPDINLTGI 176

RESULT 10  
US-10-414-378-34  
; Sequence 34, Application US/10414378  
; Publication No. US20030165981A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: 058769  
; CURRENT APPLICATION NUMBER: US/10/414,378  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: US/09/529,063  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-414-378-34

Query Match 35.4%; Score 520; DB 12; Length 185;  
Best Local Similarity 62.1%; Pred. No. 6,1e-40;  
Matches 110; Conservative 15; Mismatches 50; Indels 2; Gaps 2;  
QY 1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCYTLPTSGTLVPMCKGKPCPW 60  
DB 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAVGONAYLPFCFYTPAAGNLVPCWKGKACPV 60

DB 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAVGONAYLPFCFYTPAAGNLVPCWKGKACPV 60  
QY 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120  
DB 61 PECGNVLRTERDENVYV-TSRVWLNDFPKGDVSLITENVTLADSGIYCCRIQIPGLMN 119  
QY 121 DKLELKLIDIKAAVTPAQTAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKI 176  
DB 120 DEKNFLKLVKPAKVPAPLTQRDFTAFPRMLTTRGHGPAETQTLGSLPDINLTGI 176

RESULT 11  
US-09-966-546-22  
; Sequence 22, Application US/09966546  
; Patent No. US20020168716A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine  
; APPLICANT: Shimkets, Richard A.  
; TITLE OF INVENTION: No. US20020168716A1el Human Proteins and Polynucleotides  
; FILE REFERENCE: Cura-46 (15966-546)  
; CURRENT APPLICATION NUMBER: US/09/966,546  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 09/544,511  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-966-546-22

Query Match 29.3%; Score 430.5; DB 10; Length 142;  
Best Local Similarity 63.2%; Pred. No. 7,4e-32;  
Matches 86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 1 MFSGLTNCVLLQLLLARSLSDGYKVEGKNAVLPSCYTLPTSGTLVPMCKGKPCPW 60  
DB 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAVGONAYLPFCFYTPAAGNLVPCWKGKACPV 60  
QY 61 SOCTNELRTDERNVYQKSSRYOLKGDVSLIKVNTLDDHGTCCRIOPGLMN 120  
DB 61 PECGNVLRTERDENVYV-TSRVWLNDFPKGDVSLITENVTLADSGIYCCRIQIPGLMN 119  
QY 121 DKLELKLIDIKAAVTP 136  
DB 120 DEKNFLKLVKPGEMT 135

RESULT 12  
US-09-966-545-22  
; Sequence 22, Application US/09966545  
; Patent No. US20020172999A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine  
; APPLICANT: Shimkets, Richard A.  
; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encoding  
; FILE REFERENCE: Cura-46 (15966-546)  
; CURRENT APPLICATION NUMBER: US/09/966,545  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 09/544,511  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-966-545-22



```

; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 146
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-146

```

```

Query Match      29.3%; Score 430.5; DB 16; Length 142;
Best Local Similarity 63.2%; Pred. No. 7.4e-32;
Matches 86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 1 MFSGLTNCVLLIQLILARSLEDGYKVEGKNAYLPCSYTLPTSGTLVPMCMGKGFPCPW 60
Db 1 MFSHLPPFCVLLIILLLILTRSEVEYRREVGONAYLPCFTYTPAAGNLYFVCMGKACPV 60

QY 61 SOCTNEILRTBRNVTYKSSRYQLKGPLNGDVSLLIKNVTLDHGTGCCRIOPGLMN 120
Db 61 FECGNVVLRTDERDVNW-TSRYWLNGDFRKGDVSLTIENTVLADSGIYCCRIQIPGIMN 119

QY 121 DKLEIKLIDIKAKVT 136
Db 120 DEKFNKLVIKPEWT 135

```

Search completed: November 22, 2003, 05:56:14  
 Job time : 44.0052 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd..

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:17:26 ; Search time 23.1753 Seconds  
(without alignments)

1166.045 million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468

Sequence: 1 MFSGLTNCTVLLQLLLAR.....VVEVENSENYYCYVNSQPS 281

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267.5	18.2	451	2 S71754	cellular hepatitis
2	136.5	9.3	365	2 S71780	coxsaackie- and ade
3	131	8.9	503	2 UC5287	SHP substrate-1 pr
4	123	8.4	309	2 I49503	B-lymphocyte activ
5	120.5	8.2	403	2 I52590	m3-B isoform - mo
6	118.5	8.1	538	2 JC2457	vascular cell adhe
7	116	7.9	459	2 A46254	CD4 precursor - ra
8	111.5	7.6	247	2 A55717	myelin/oligodendro
9	106	7.2	186	2 I61783	sodium channel bet
10	106	7.2	408	1 LKRT2	sodium channel bet
11	105	7.1	215	2 A57843	proteoglycan link
12	104	7.1	354	1 LKHU	proteoglycan link
13	104	7.1	509	2 UC5288	SHP substrate-1 pr
14	104	7.1	513	2 UC5289	SHP substrate-1 pr
15	103.5	7.1	354	1 S42938	proteoglycan link
16	102.5	7.0	457	2 A27449	T-cell surface gly
17	102	6.9	457	1 RWMST4	T-cell surface gly
18	102	6.9	457	1 T16525	hypothetical prote
19	101.5	6.9	458	1 RWHUT4	T-cell surface gly
20	101.5	6.9	1021	2 I39207	leukocyte surface
21	101.5	6.9	1323	2 PNO568	connectin 3B - chi
22	101.5	6.9	4162	2 T42633	connectin/citlin -
23	101	6.9	458	2 JC1509	biliary glycoprote
24	101	6.9	521	2 S34338	biliary glycoprote
25	100	6.8	364	2 A30521	myeloid cell surfa
26	98.5	6.7	346	2 S46993	elk ligand - human
27	98.5	6.7	354	1 S04243	proteoglycan link
28	98	6.7	562	2 G02426	interleukin-1 rece
29	97	6.6	3295	2 AE0074	probable adhesin y

30	96.5	6.6	1006	2 T31685	beta-galactosidase
31	96	6.5	345	2 I48780	Srral/Ep192 protei
32	95.5	6.5	349	2 A34815	carcinoembryonic a
33	95.5	6.5	1199	2 T23005	hypothetical prote
34	95	6.5	246	2 A47712	myelin/oligodendro
35	95	6.5	368	2 S17428	interleukin-1 rece
36	95	6.5	764	1 ORHUG5	secretory componen
37	95	6.5	967	2 S66852	hypothetical prote
38	95	6.5	1259	2 S36126	neural cell adhest
39	94.5	6.5	1033	2 S19247	cell adhesion prot
40	94	6.4	272	2 C49209	outer surface prot
41	94	6.4	458	1 WMSR1	biliary glycoprote
42	94	6.4	521	2 JC1508	biliary glycoprote
43	94	6.4	523	2 I50478	neuroilin - goldfis
44	94	6.4	1257	1 A41060	neural cell adhest
45	92.5	6.3	218	2 B47712	myelin/oligodendro

#### ALIGNMENTS

##### RESULT 1

S71754 cellular hepatitis A receptor HAVcr-1 precursor - green monkey

N:Alternate names: surface glycoprotein

C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 05-Jun-1998

C:Accession: S71754

R:Kaplan, G.; Totenuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstein, S.M.

EMBO J. 15, 4282-4296, 1996

A:Title: Identification of a surface glycoprotein on African green monkey kidney cells a

A:Reference number: S71754; MUID:97015129; PMID:8661957

A:Accession: S71754

A:Molecule type: mRNA

A:Residues: 1-451 <Kap>

A:Cross-references: EMBL:X98252; NID:g1526573; PID:e247449; PID:g1526574

A:Experimental source: kidney

C:Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <Mat>

Query Match	18.2%	Score 267.5;	DB 2;	Length 451;
Best Local Similarity	21.7%;	Pred. No. 1.86-14;		
Matches	97;	Conservative 51;	Mismatches 96;	Indels 203; Gaps 11;
QY	7 LNCVLLLOLLARSLEDGKVE--VGKNAVLPESYTLPTSGLVPMCKGKGFPMQOCT 64			
DB	3 LQVILSLILHLADSVADVVDGAGLSITLPCRY----NGAITSMCNMGTSVFSQP 58			
QY	65 NEILRTDERNWTYQKSRYOLKGDINKGVDLSLIKNTVLDHGYCCRIOPGLMNDKL 124			
DB	59 DGIWITGTHVTRKTRKRLKGLNLSRVDLSLTIANAVSDSGLYCRAKSGFNDWKI 118			
QY	125 ELKDIRAKVT----- 136			
DB	119 TISLKIQRPTVIVATVATSTVPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTL 178			
QY	137 ----- 136			
DB	179 LPTTTPPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTL 238			
QY	137 ----- 136			
DB	239 TTTTLPPTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 298			
QY	137 -----PAQT-----AHGDSYASP-RLLTERNGSEFQTLVTLANNNGKIST 178			
DB	299 PVATSPSPQAEKHPVTLGATRTQPTSPSLVYTTDGSPTVRESSDGLWNNNOTQLSP 358			
QY	179 WADIKDSGETIR--AIHIGVGSAGTLALIGVLLKRYSKKKLSLSLTLANL 236			
DB	359 -----EHSPOVNTTBGIYAGVCISV-LVLAVIGVIAKKYFP-KKEIQQLS-VSPSNH 410			

QY 237 PPGGLANAGAVRIRSEENITYTEENY 263  
 Db 411 QPKTLQNAVKEVHADNITY-IENNL 436

## RESULT 2

coxackie- and adenovirus receptor - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 02-Apr-2002  
 C:Accession: J07780

R:Thoclem, I.; Keyaerts, E.; Lindberg, M.; Van Ranst, M.  
 Biochem. Biophys. Res. Commun. 288, 805-808, 2001

A:Title: Characterization of a cDNA encoding the bovine coxackie and adenovirus receptor  
 A:Reference number: J07780

A:Contents: Liver  
 A:Accession: J07780  
 A:Molecule type: mRNA  
 A:Residues: 1-365 <THO>  
 A:Cross-references: GB:AY033651  
 C:Comment: This protein serves as the primary adenoviral attachment site on bovine cells

Query Match 9.3%; Score 136.5; DB 2; Length 365;  
 Best Local Similarity 22.5%; Pred. No. 0.0011;  
 Matches 61; Conservative 44; Mismatches 99; Indels 67; Gaps 11;

QY 12 LLLQLLARSLEB--GYRVEV-----GKNATLPSCSYLPTSGTIVPMCKGKFCF 59  
 Db 3 LLLRFLLCGVADFTGRLITTPPEQMIERAKGTATVLPCKFTLG-----PEDQGLDIE 56  
 QY 60 WISQCTELLRTBERNVTYQKSSRY-----QLKG-----DLNKGDVSLIKNVTLDHG 107  
 Db 57 WLLSPADNOKVDVILLYSGDKTYDYODLKGKRVHFTSNLKGSDASINVTNLQLSDIG 116  
 QY 108 TYCCRI-QPPGLMNDKLEKLDIKAA-----KVTPAQTA---HGD 144  
 Db 117 TYQCKYKKAQVGN-KKIQTLVLRKSGRCYVDGSEBIGNPKLCEKESGLPIRYEM 175  
 QY 145 STTASRITLTERNGSETQTLVTLHNNNGTKISTWADKIDGSETRTIHHGV----- 198  
 Db 176 QKLSQDKKPTSWLPMTSPVISVKNASAEVSGTCTYCVNRVGSDOCLRLDVPVPSNR 235  
 QY 199 ----GVSAGLTALITGVILKWSCKKK 224  
 Db 236 AGTIAGAVIGTLALVALIV--FCHKKR 264

## RESULT 3

SHP substrate-1 protein - human

C:Species: Homo sapiens (man)  
 C:Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000  
 C:Accession: J05287

R:Yamato, T.; Matsuoka, T.; Matsuda, Y.; Takahashi, N.; Ochi, F.; Fujioka, Y.;  
 Biochem. Biophys. Res. Commun. 231, 61-67, 1997  
 A:Title: Mouse and human SHP-1: Molecular cloning of cDNAs and chromosomal localization  
 A:Reference number: J05287; MUID:97223399; PMID:9070220

A:Contents: Brain  
 A:Accession: J05287  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-503 <YAM>  
 A:Cross-references: DDBJ:D86043; NID:g1864010; PINN:BA12974.1; PID:g1864011  
 C:Comment: This protein is a glycosylated receptor-like protein and plays a role in cell  
 C:Genetics: acts as a docking protein and induce translocation of SHP-2 from the cytosol to the pl  
 A:Gene: shps-1  
 A:Map position: 20p13

Query Match 8.9%; Score 131; DB 2; Length 503;  
 Best Local Similarity 22.5%; Pred. No. 0.0047;  
 Matches 67; Conservative 45; Mismatches 124; Indels 62; Gaps 13;

QY 13 LLLQLLARSLEB-----EDGYKVEGKNAVLPSCY--LPTSGTLVPMCKG 54  
 Db 14 LLLQLLARSLEB-----EDGYKVEGKNAVLPSCY--LPTSGTLVPMCKG 54  
 QY 55 KGFPCWQCTNELRTDE-----RNATYQKSSRYQKLDKNGDVSLIKNVTLDHGT 109  
 Db 70 RQAGP-----ARLITNQGKGFPRVITYSESKRE-----NMDSISISNITPADGTY 119  
 QY 110 CCRIFPPGLMNDKLE---LKLDIKAAKVTPTAHDSTTASPR---TLTERNGSET 162  
 Db 120 YC-VKFRKGSPTDEKSGAGTELIVRAKPSAP--VSGPAPARATQHTVTSFCESHGSP 176  
 QY 163 QTLVTLHNNNGTKISTWADKIDGSETRTIHHGVSAAGTLTLLIIGVILKRYSCCK 222  
 Db 177 RDTLKWRFGNGLSDFTQTNVDPVGSVSYSTHSTAKV-----VLTREDVHS 223  
 QY 223 KLLSSLSLITLANLPGLANAGAVRIRSEENITYTEENYEVENSNEYCVNSQOP 280  
 Db 224 QVICAVAVTLQGDRLRGTALNLSER-IVPPTLEVTQGPV-RAENQVAVTCQVRKFP 279

## RESULT 4

149503  
 B-lymphocyte activation antigen 7 precursor - mouse

N:Alternate names: MB7-2  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
 C:Accession: U9503; S17291; I49521  
 R:Selyakumar, A.; White, P.C.; Dupont, B.  
 Immunogenetics 38, 292-295, 1993

A:Title: Genomic organization of the murine B-lymphocyte activation antigen B7.  
 A:Reference number: U9503; MUID:93307789; PMID:768531  
 A:Accession: U9503

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-309 <RES>  
 A:Cross-references: GB:U12589; NID:g293299; PIDN:AA37240.1; PID:g293301

R:Freeman, G.J.; Gray, G.S.; Gimmi, C.D.; Lombard, D.B.; Zhou, L.O.; White, M.; Fingerot  
 J. Exp. Med. 174, 625-631, 1991

A:Title: Structure, expression, and T cell costimulatory activity of the murine homolog  
 A:Reference number: S17291; MUID:91341422; PMID:1714935

A:Accession: S17291

A:Molecule type: mRNA

A:Residues: 1-274, 'R', 279-309 <PRE>  
 A:Cross-references: EMBL:X60958; NID:g50111; PIDN:CAA4291.1; PID:g50112

R:Inobe, M.; Linsley, P.S.; Ledbetter, J.A.; Nagai, Y.; Tanakoshi, M.; Ueda, T.  
 Biochem. Biophys. Res. Commun. 200, 443-449, 1994

A:Title: Identification of an alternatively spliced form of the murine homologue of B7.  
 A:Reference number: U9521; MUID:94220123; PMID:7513163

A:Accession: U9521

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-143, 238-274, 'R', 279-309 <RE2>  
 A:Cross-references: GB:D16220; NID:g505118; PIDN:BA03748.1; PID:g994769

C:Genetics:

A:Gene: B7

A:Introns: 37/1, 143/1, 237/1, 275/1  
 A:Superfamily: B-lymphocyte restricted antigen B7

C:Keywords: alternative splicing

Query Match 8.4%; Score 123; DB 2; Length 309;  
 Best Local Similarity 24.2%; Pred. No. 0.012; Pred. No. 0.012;  
 Matches 67; Conservative 31; Mismatches 89; Indels 90; Gaps 12;  
 QY 11 LLLQLLARSLEB-----SLDEGYKVEGKNAVLPSCY--LPTSGTLVPMCKG 51  
 Db 21 LLLRFLLCGVADFTGRLITTPPEQMIERAKGTATVLPCKFTLG-----PEDQGLDIE 56  
 QY 52 KGFPCWQCTNELRTDEBNVTYQKSSRYQKLDKNGDVSLIKNVTLDHGTGCC 111  
 Db 81 IAGK-LKWPVEYKN--RTLYDNTTY-----SLTIIGLVLSDRGTYSC 119  
 QY 112 RIQFP--GLMNDKLEL-KLDIKAAKVTPTAHDSTTASPR----- 151



Db 120 VQKKERGTVEVKHLALVKLSIKADFSTPNITESGNBPADTKRITCFASGGFPKPRFSL 179  
Qy 152 -----TLTERNGSEFQTLVTLHNNNGTK-----IS---TWABEIK 184  
Db 180 ENGELPOINTNTTISQDPESELTYTSSQDPEFNTNHTIKLICYGDAHVSDFTEKPE 239  
Qy 185 DSGETIRTAIHIGVVSAGLTALILIGVILKMYSC 221  
Db 240 DPEKNTLVLFAGAGFAGVITVVI--VVIKCF-CK 273

## RESULT 5

152590  
m33-B isoform - mouse  
C:Species: Mus sp. (mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: 152590  
R:Chilian, E.Z.; Beverley, P.C.; Young, B.D.; Watt, S.M.  
Blood 83, 3188-3198, 1994  
A:Title: Molecular cloning of two isoforms of the murine homolog of the myeloid CD33 and  
A:Reference number: 152590, MUID:94250900, PMID:8193354  
A:Accession: 152590  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-403 <RES>  
A:Cross-references: GB:571345, NID:9551352, PIDN:AA830842.1, PID:9551353

Query Match 8.2%; Score 120.5; DB 2; Length 403;  
Best Local Similarity 23.7%; Pred. No. 0.026;  
Matches 65; Conservative 32; Mismatches 116; Indels 61; Gaps 8;

Qy 10 VLLLLQLLLARSL-----DGYVEVGNKMYLPCSYLP-----TSGTLVPMCMGKGRCP 59  
Db 7 LFLCAGSLADLEFQVAPSVTEBSLCVHPSPVFPYSIKLTLGVTGSMWRKV-- 64  
Qy 60 WSQCTNELLRTDERNVTYQKSS--RYQKGLINKGVSLLIKNTVLDHGTGCRIOFP 117  
Db 65 -SLHEDSPVATSDRQLVQKATQGRFQLGDPQKHDCSLFRDQKNDTGMVFRVVRP 123  
Qy 118 LM-NDKLELKLIDIKAAKTPAQTAGDSTTASPRITL----- 154  
Db 124 FVRYSYKKSQSLVHTLSRTPDIIIPGTLGAGYPSNLTGCVPAECGPTPTFSWMSTA 183  
Qy 155 -----TERNGSEFQTLVTLHNNNGTKISTWADIKDSETRTAIHIGVVSAG----- 203  
Db 184 LTSLSSTRTDSSVLTFTPOPDHGTCL-TCLVTSAGAVTERTIQLNVTKRSQOMREL 242  
Qy 204 -----LTLALILIGVILKMYSCKKKLS 226  
Db 243 LVAVGEATVKLLILGLCLVFLIWMFCRRKTKLS 276

## RESULT 6

JC2457  
vascular cell adhesion protein - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 08-Oct-1999  
C:Accession: JC2457  
R:Tsang, Y.T.M.; Haskard, D.O.; Robinson, M.K.  
Biochem. Biophys. Res. Commun. 201, 805-812, 1994  
A:Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule.  
A:Reference number: JC2457, MUID:94271236, PMID:7516159  
A:Accession: JC2457  
A:Molecule type: mRNA  
A:Residues: 1-538 <TSA>  
A:Cross-references: EMBL:U08351, NID:9474382, PIDN:AAA21542.1, PID:9474383  
C:Keywords: glycoprotein, transmembrane protein  
F:497-517/Domain: transmembrane #stratus predicted <TM>  
F:75,157,271,330,360/binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 8.1%; Score 118.5; DB 2; Length 538;  
Best Local Similarity 25.1%; Pred. No. 0.056;

Matches 63; Conservative 42; Mismatches 81; Indels 65; Gaps 15;  
Qy 63 CTNELRTDERNVTYQKSSRYQKGLINKGVSLLIKNTV-----DDHGTGCRIOF 115  
Db 244 CTSEGLPAPD--ISWSK-----LDNGDQQLSGNATLTLANMEBSGIYECGVN 293  
Qy 116 PGLMNDKLELKLID-----AAKVTAPAQTAGDSTT-----ASPRITLITE--RNGS 160  
Db 294 PVGTNRKEVELTQVAPRDTTISVNPSTLEBSSVVMTCSSDGFAPKILMSKLDGN 353  
Qy 161 ETQTLVTLHNNNGTKISTWADIKDSETRTAIHIGVVS-AGLT--LALITGV----- 212  
Db 354 -----LEPLSEN--TTLTLTSTKMDSG-----IYVCEGINOAGINKKEVELIIQAAPKD 401  
Qy 213 LILKMYSCKKKLSLSLI--TLANLPPGL-----ANAGAVRIRSENYTYT-----E 259  
Db 402 LQLPAPSESVKEDTVITCTGAVPFLIILKKKAEFDVTKSDGATTHIRARLAD 461  
Qy 260 ENVEVENSNE 270  
Db 462 AGVVECESKNE 472

## RESULT 7

A46254  
CD4 precursor - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
C:Accession: A46254  
R:Hague, B.F.; Sawadkhol, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
A:Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi  
A:Reference number: A46254, MUID:92390370, PMID:1518821  
A:Accession: A46254  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-459 <HAG>  
A:Cross-references: GB:M92840, NID:g164871, PIDN:AAA31198.1, PID:g164872  
C:Note: sequence extracted from NCBI backbone (NCBI:112732, NCBI:P:112733)  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
F:322-372/Domain: immunoglobulin homology <IM>

Query Match 7.9%; Score 116; DB 2; Length 459;  
Best Local Similarity 26.8%; Pred. No. 0.074;  
Matches 45; Conservative 21; Mismatches 54; Indels 48; Gaps 7;

Qy 9 CVLLILQLLLARSLLEDGYKVEVGNAYLPCSYLPTSGTLVPMCMGKGFPMSCQTNEL 68  
Db 9 CLILVLPALIPATWGTVVRGK-----AGALVELP-----CQSQKRNVSF 51  
Qy 69 RTDERNVTY-----QKSSRYQKGD-----LNKGVSLIKNTVLDHGTG 109  
Db 52 NKRHANOVKILGNQSSSSSFWMKGNPLSNRVESKKNMWDQGFPLVTDLRMDSGTY 111  
Qy 110 CCRIOFPGLMNDKLELKLIDIKAAKVP-AQTAGDSTTASPRITLITE 156  
Db 112 ICEV-----GDKMEVELVFRITANPNRLHAGSL-----TLTLE 148

## RESULT 8

A55717  
myelin/oligodendrocyte glycoprotein precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 08-Oct-1999  
C:Accession: A55717, C47712  
R:Daubae, P.; Pham-Dinh, D.; Dautigny, A.  
Genomics 23, 36-41, 1994  
A:Title: Structure and polymorphism of the mouse myelin/oligodendrocyte glycoprotein gen  
A:Reference number: A55717, MUID:95130110, PMID:7829100  
A:Accession: A55717  
A:Molecule type: DNA  
A:Residues: 1-247 <DAU>  
A:Cross-references: GB:L29498

R.Pharm.Dinh, D.; Mattei, M.G.; Nusbaum, J.L.; Roussel, G.; Pontarotti, P.; Roeckel, N.;  
Proc. Natl. Acad. Sci. U.S.A. 90, 7990-7994, 1993  
A>Title: Myelin/Oligodendrocyte glycoprotein is a member of a subset of the immunoglobul  
A:Reference number: A47712; MUID:93376728; PMID:8367453  
A:Accession: C47712  
A:Molecule type: mRNA  
A:Residues: 30-95, 'E', 97-247 <PHA>  
A:Cross-references: GB:L20942; NID:g399588; PIDN:AAA03180.1; PID:g399589  
C:Genetics:  
A:Gene: MOG  
A:Map position: 17  
A>Note: encoded within the MHC  
C:Function:  
A:Description: may be involved in lipid interaction; may be involved in cell-cell commu  
C:Keywords: glycoprotein; myelin; transmembrane protein  
F:1-23/Domain: signal sequence #status predicted <Sig>  
F:30-247/Product: myelin/oligodendrocyte glycoprotein #status predicted <MAT>  
F:151-179/Domain: transmembrane #status predicted <TM1>  
F:204-229/Domain: transmembrane #status predicted <TM2>  
F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.6%; Score 111.5; DB 2; Length 247;  
Best Local Similarity 26.6%; Pred. No. 0.08;  
Matches 37; Conservative 28; Mismatches 53; Indels 21; Gaps 5;

QY 8 NCYLLLLQLLA-----RSLDGYKE--VGNAYLPCTYTLPTSGTIVPMCKGK 55  
DB 11 SCPLSLLLLLQLLQSCSYAGQFVIGPGYPIRALVDEDELPRISPGKNAKGMVEGWYR 70  
QY 56 GFCFMSQCTNELRTDERANTYOKSSRYQ----LKGDINKGVSLLIKNVTLDHGTGTC 110  
DB 71 S--PFRVYH--LYRNGKQDAEQAPERYKGRTELLKETTISEGKVTIRIQNVRSDSGYT 126  
QY 111 CRIQPGPLMNDKKLLEKLD 129  
DB 127 CFEFRDHSYQBEAMLEKVE 145

RESULT 9  
161783  
sodium channel beta 2 subunit - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I61783  
R:Jenson, L.L.; Regsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Schuer,  
Cell 83, 433-442, 1995  
A>Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Trans  
A:Reference number: A57843; MUID:96067641; PMID:8521473  
A:Accession: I61783  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-186 <RES>  
A:Cross-references: EMBL:U37147; NID:g1086498; PIDN:AAB60506.1; PID:g1086499  
C:Genetics:  
A:Gene: SCN2B  
A:introns: 50/3; 120/2

Query Match 7.2%; Score 106; DB 2; Length 186;  
Best Local Similarity 23.2%; Pred. No. 0.16;  
Matches 52; Conservative 29; Mismatches 77; Indels 66; Gaps 8;

QY 31 GKNAYLPCSYT--LPTSGTIVPMCKGKFCPMSCQCTNELRTDERANTYOKSSRY----Q 84  
DB 14 GSDTRLPCFPNSCYTVNHHKPSLMTYQEC--SNCEENVLQFRMKIINLKERFQDRVE 71  
QY 85 LKGDINKGVSLLIKNVTLDHGTGTCRIQFGLMNDKKLELTDIKAAKVTPTAQTADHD 144  
DB 72 FSGNPKYVSVTLKRVQLEDEGIYNCYITNPDRRHGKIKLYQV--LLEPPER----D 126  
QY 145 STTASPRLLTTERNGSETQTLVTLHHNNNGTKISTMADEIKDGETITRAIHIGVGSAGI 204  
DB 127 STVA-----VIVGASVGG 139

QY 205 TIALITGVILKWSYSC---KKKKLSLSLITLNLPPGGLANA 244  
DB 140 FLAAVVI--LVLMVXCVRRKKEQKLSLTDLKTREEGKTGDEGNA 181

RESULT 10  
LKRT2  
proteoglycan link protein 2 precursor - rat  
N:Alternate names: cartilage link protein  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 13-Aug-1986 #sequence\_revision 24-Oct-1997 #text\_change 13-Nov-1998  
C:Accession: A28654; A24880; A02869  
R:Rhodes, C.; Doege, K.; Sasaki, M.; Yamada, Y.  
J. Biol. Chem. 263, 6063-6067, 1988  
A>Title: Alternative splicing generates two different mRNA species for rat link protein.  
A:Reference number: A28654; MUID:88198139; PMID:2452158  
A:Accession: A28654  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-408 <RHO>  
R:Doege, K.; Hassell, J.R.; Caterston, B.; Yamada, Y.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3761-3765, 1986  
A>Title: Link protein cDNA sequence reveals a tandemly repeated protein structu  
A:Reference number: A24880; MUID:86233314; PMID:3459153  
A:Accession: A24880  
A:Molecule type: mRNA  
A:Residues: 180-408 <DOE>  
R:Neame, P.D.; Christner, J.E.; Baker, J.R.  
J. Biol. Chem. 261, 3519-3535, 1986  
A>Title: The primary structure of link protein from rat chondrosarcoma proteoglycan aggr  
A:Reference number: A02869; MUID:86140139; PMID:2419334  
A:Accession: A02869  
A:Molecule type: protein  
A:Residues: 16-33, 'A', 89-375, 'W', 377-408 <NEA>  
C:Comment: This protein was extracted from rat chondrosarcoma.  
C:Comment: Residues 259-277 and 358-376 (approximately) form epitopes for the species-no  
C:Comment: Link proteins interact with and stabilize aggregates of hyaluronic acid and c  
C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat  
C:Keywords: cartilage; chondroitin sulfate proteoglycan; duplication; glycoprotein; hyal  
F:1-15/Domain: signal sequence #status predicted <Sig>  
F:16-408/Product: proteoglycan link protein 2 #status experimental <MAT>  
F:108-195/Domain: immunoglobulin homology <IMM>  
F:230-307/Domain: link protein repeat homology <LNK1>  
F:328-404/Domain: link protein repeat homology <LNK2>  
F:110/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:115-193, 235-306, 259-280, 333-403, 358-379/Disulfide bonds: #status experimental

Query Match 7.2%; Score 106; DB 1; Length 408;  
Best Local Similarity 29.9%; Pred. No. 0.43;  
Matches 40; Conservative 21; Mismatches 39; Indels 34; Gaps 8;

QY 20 RSLDGYKEV-----GKNAYLPCSYTLPTSGTIVPMCKGKF---CPWSQCT 64  
DB 86 RTTENGPRLLVFAEQAKVFSHRGQVNTLPCKRYRD-----PTAGSGIHKIRIKWTKLT 139  
QY 65 NELRTDERANTY--YOKSS-----RYQLKGDLNKGVSLTIKNVTLDHGTGTCRIQFP 116  
DB 140 SYLREVDVDFVSMGHHKTYGGYGVGFVFLKGG--SDNDASLITLTDLBDYGRYKCEV--IB 197  
QY 117 GLMNDK--KLELK 127  
DB 198 GLEDTAVVALELQ 211

RESULT 11  
A57843  
sodium channel beta 2 subunit - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: A57843  
R:Jenson, L.L.; Regsdale, D.S.; DeJongh, K.S.; Westebroek, R.E.; Reber, B.F.X.; Schuer,  
Cell 83, 433-442, 1995  
A>Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Trans

A:Reference number: A57843; MUID:96067641; PMID:8521473  
A:Accession: A57843  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1215 <RES>  
A:Cross-references: EMBL:U37026; NID:g1086496; PIDN:AAC53967.1; PID:g1086497  
C:Genetics:  
A:Gene: SCNB2

Query Match 7.2%; Score 105; DB 2; Length 215;  
Best Local Similarity 23.2%; Pred. No. 0.23;  
Matches 52; Conservative 28; Mismatches 78; Indels 66; Gaps 8;

31 GKNAVPCSYT--LPTSGTLVPMCMGKGFPCMSQCTNELRTDERNTYOKSSRY---Q 84  
43 GSTRLPCTNSCYVHHKPSLMTWYQEC--SNCSSEMLQPMKINKLERFGDRI 100  
85 LKCDLMDKGVSLIIRKNTLDHGTCCRIQFPGMLNDKLELKDIAKAVTPAQTAGD 144  
101 FSGNPSKYDVSVTLKNVQLEDEGIYNCYITNPDRHGHGKIYLVQ--LLEVPPEP---D 155  
145 STASRRTLTENSGSTQTLVTLHNNNGKISRWAEIIDSGETITTAHIGVSAQL 204  
156 STVA-----VIVGASVVG 168

205 TLLAIGVLLIKWYSC---KKKLSLSLITLANLPGLANA 244  
169 FLAVVI--LVLMVVKCYRRKKEQKLTSDDKTEBEKGTDSEGN 210

RESULT 12  
LKHU  
proteoglycan link protein precursor [validated] - human  
N:Alternate names: cartilage link protein  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 08-Dec-2000  
C:Accession: S14914; S04244; S03868; A36308; S14926; S09309; S08041  
R:Dudhia, J.; Hardingham, T.E.  
Nucleic Acids Res. 18, 1292, 1990  
A>Title: The primary structure of human cartilage link protein.  
A:Reference number: S14914; MUID:90206798; PMID:2320422  
A:Accession: S14914  
A:Molecule type: mRNA  
A:Residues: 1-354 <DUD>  
A:Cross-references: EMBL:X17405; NID:9463246; PIDN:CAA35462.1; PID:g34378  
A>Note: the authors translated the codon GAT for residue 93 as Ala  
R:Dudhia, J.; Hardingham, T.E.  
J. Mol. Biol. 206, 749-753, 1989  
A>Title: Appendix. Isolation and sequence of cDNA clones for pig and human cartilage link protein.  
A:Reference number: S04243  
A:Accession: S04243  
A:Molecule type: mRNA  
A:Residues: 223-354 <DUD2>  
A:Cross-references: EMBL:Y00166  
A>Note: the authors translated the codon CTG for residue 264 as Arg and TTG for residue 265.  
R:Nguyen, Q.; Murphy, G.; Roughley, P.J.; Mort, J.S.  
Biochem. J. 259, 61-67, 1989  
A>Title: Degradation of proteoglycan aggregate by a cartilage metalloproteinase. Evidence  
A:Reference number: S03868; MUID:89246328; PMID:2719651  
A:Accession: S03868  
A:Molecule type: protein  
A:Residues: 16-35 <NGU>  
R:Osborne-Lawrence, S.L.; Sinclair, A.K.; Hicks, R.C.; Lacey, S.W.; Eddy Jr., R.L.; Byer  
Genomics 9, 562-567, 1990  
A>Title: Complete amino acid sequence of human cartilage link protein (CRTLI) deduced from  
A:Reference number: A36308; MUID:91139126; PMID:2286376  
A:Accession: A36308  
A:Status: preliminary; nucleic acid sequence not shown, not compared with conceptual tra  
A:Molecule type: mRNA  
A:Residues: 1-354 <OSB>  
R:Dudhia, J.; Hardingham, T.E.  
Nucleic Acids Res. 18, 2214, 1990  
A>Title: The primary structure of human cartilage link protein.

A:Reference number: S14926; MUID:90245703; PMID:2336413  
A:Accession: S14926  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92, 'A', 94-354 <NUC>  
R:Nguyen, Q.; Mort, J.S.; Roughley, P.J.  
Biochem. J. 266, 569-573, 1990  
A>Title: Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L than  
A:Reference number: S09309; MUID:90197639; PMID:2317204  
A:Accession: S09309  
A:Molecule type: protein  
A:Residues: 16-38, 'X', 40-55, 57-60, 62-65, 'X', 67, 'X', 69-80 <NG2>  
C:Genetics:  
A:Gene: GDB:CRTL1  
A:Cross-references: GDB:125232; OMIM:115435  
A:Map position: 5q13-5q14  
C:Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat  
C:Keywords: cartilage; duplication; extracellular matrix; glycoprotein  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-354/Product: proteoglycan link protein #status experimental <MAT>  
F:54-141/Domain: immunoglobulin homology <IMM>  
F:176-253/Domain: link protein repeat homology <LNK1>  
F:274-350/Domain: link protein repeat homology <LNK2>  
F:21,56/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:61-139/Disulfide bonds: #status predicted

Query Match 7.1%; Score 104; DB 1; Length 354;  
Best Local Similarity 24.2%; Pred. No. 0.53;  
Matches 39; Conservative 27; Mismatches 47; Indels 48; Gaps 8;

10 VLLLLQLLARSLEDEYKVE-----VGNNAVPCSYTL 42  
6 LVLVISICWADHLSNDVTLDDHRAIHQAENGPHLVEAQAVFHSRGNAVTLPCFPR 65  
43 PTSGTLVPMCMGKGF---CPMSQCTNELRTDERNT--YOKSS-----RYLKGDLN 90  
66 D-----PFAFGSGIHRIKIRIKMKLTSDYLKEVDVFWSMGYHKKTYGGYGRVFLKGG-S 118  
91 KGPVSLIRKNTLDHGTCCRIQFPGMLNDKLELKDIAKAVTPAQTAGD 131  
119 DSDASLVITDLTLEDYGRKCEV-IRGL-EDTVVVALDQ 157

RESULT 13  
JCS288  
SHP substrate-1 protein, 509 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000  
C:Accession: JCS288  
R:Yamao, T.; Matczak, T.; Amano, K.; Matsuda, Y.; Takahashi, N.; Ochi, F.; Fujioaka, Y.;  
Biochem. Biophys. Res. Commun. 231, 61-67, 1997  
A>Title: Mouse and human SHP-1: Molecular cloning of cDNAs and chromosomal localization  
A:Reference number: JCS287; MUID:97223399; PMID:9070220  
A:Accession: JCS288  
A:Molecule type: mRNA  
A:Residues: 1-509 <YAM>  
A:Cross-references: DDBJ:D87967; NID:g1864012; PIDN:BAI3520.1; PID:g1864013  
C:Comment: This protein is a glycosylated receptor-like protein and plays a role in cell  
acts as a docking protein and induce translocation of SHP-2 from the cytosol to the plasma  
C:Genetics:  
A:Gene: shps-1  
A:Map position: 2

Query Match 7.1%; Score 104; DB 2; Length 509;  
Best Local Similarity 20.4%; Pred. No. 0.84;  
Matches 68; Conservative 39; Mismatches 97; Indels 130; Gaps 14;

12 LLLQLLARSLE-----EDGKVEGKNAVPCSYT--LPTSGTLVPMCMG 54  
14 LLLCLLSASCFCTGVTKGLKVTQPEKSVSAAGDSTVLNCTLTSLPVG---PLKMY 69  
55 KGFCPMSQCTNELRTDERNTYOKSSRYOLKGD-----LNKGVSLIIRKNV 101

```

Db      70 RG-----VQSRLIYSFTGEHPPRTVNSDATKRNNDPSIRISNV 111
Qy      102 TLDDHGYTCRIQFPGMLNDKLEKLD-----IKAKVTPAQT----- 140
Db      112 TPEDAGTYVC-VKFGKGPSEPDTEIQSGGSTEYVLAKSPPEVSGPADRGIPDQKVNPT 170
Qy      141 --AHGDSSTTASPRITLITE--RNGSETQTLVTLHNNNGTKISTWABEIKDSGETIRTAIHI 196
Db      171 CKSHG----FSPRNTTLKWKFGQGEIHLHLETTVNSGKNVS-----YNISSTVR----- 215
Qy      197 GVGVSAGLTLALIIIGVLIKWYSCKKKLSLITLANLPQGLAN-----AGAVRI 249
Db      216 -----VLNNSMDVHSKVICEVAHITLDRSPLRGIANLSNFIKVSPTVKV 259
Qy      250 RSE-----ENIYTIERNVYEVSNSN 269
Db      260 TQQSPTSMNQVNLTCRAERFYEDQLIWLNGN 293

```

## RESULT 14

```

JCS289
SHP substrate-1 protein, 513 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C/Accession: JCS289
R/Yamao, T.; Maizaki, T.; Amano, K.; Matsumura, Y.; Takahashi, N.; Ochi, F.; Fujioka, Y.;
Biochem. Biophys. Res. Commun. 231, 61-67, 1997
A/Title: Mouse and human SHP-1: Molecular cloning of cDNAs and chromosomal localization
A/Reference number: JCS289; MUID:97223399; PMID:9070220
A/Contents: Bzrain
A/Accession: JCS289
A/Molecule type: mRNA
A/Residues: 1-513 <YAM>
A/Cross-references: DBJ:087968; NID:91864014; PIDN:BA13521.1; PID:91864015
C/Comment: This protein is a glycosylated receptor-like protein and plays a role in cell
act as a docking protein and induce translocation of SHP-2 from the cytosol to the pla
C/Genetics:
A/Gene: shp-1
A/Map position: 2

```

```

Query Match      7.1%; Score 104; DB 2; Length 513;
Best Local Similarity 20.4%; Pred. No. 0.85;
Matches 68; Conservative 39; Mismatches 97; Indels 130; Gaps 14;

Qy      12 LILQLLARSL-----EDGYKVEGKNAYLPQSYT--LFTSGTLVPMCWG 54
Db      14 LILCLLSASCFCTGVTKELKVTQPEKSVSAAGDSVIANCTLSILPVG---PIKMY 69
Qy      55 KGFCEWSCQTNELTDERNYVYOKSSRYQLKGD-----LNKGVSLIKNV 101
Db      70 RG-----VQSRLIYSFTGEHPPRTVNSDATKRNNDPSIRISNV 111
Qy      102 TLDDHGYTCRIQFPGMLNDKLEKLD-----IKAKVTPAQT----- 140
Db      112 TPEDAGTYVC-VKFGKGPSEPDTEIQSGGSTEYVLAKSPPEVSGPADRGIPDQKVNPT 170
Qy      141 --AHGDSSTTASPRITLITE--RNGSETQTLVTLHNNNGTKISTWABEIKDSGETIRTAIHI 196
Db      171 CKSHG----FSPRNTTLKWKFGQGEIHLHLETTVNSGKNVS-----YNISSTVR----- 215
Qy      197 GVGVSAGLTLALIIIGVLIKWYSCKKKLSLITLANLPQGLAN-----AGAVRI 249
Db      216 -----VLNNSMDVHSKVICEVAHITLDRSPLRGIANLSNFIKVSPTVKV 259
Qy      250 RSE-----ENIYTIERNVYEVSNSN 269
Db      260 TQQSPTSMNQVNLTCRAERFYEDQLIWLNGN 293

```

RESULT 15  
S42938  
proteoglycan link protein precursor - horse

```

N/Alternate names: Cartilage link protein
C/Species: Equus caballus (domestic horse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C/Accession: S42938
R/Duchia, J.; Platt, D.
submitted to the EMBL Data Library, March 1994
A/Reference number: S42938
A/Accession: S42938
A/Molecule type: mRNA
A/Residues: 1-354 <DDU>
A/Cross-references: EMBL:X78077; NID:9459438; PIDN:CA454987.1; PID:9459439
C/Superfamily: proteoglycan link protein; immunoglobulin homology; link protein repeat h
C/Keywords: cartilage; duplication; extracellular matrix; glycoprotein
F/1-15/Domain: signal sequence #status predicted <SIG>
F/16-354/Product: proteoglycan link protein #status predicted <MAT>
F/54-141/Domain: immunoglobulin homology <IMM>
F/176-253/Domain: link protein repeat homology <LINK1>
F/274-350/Domain: link protein repeat homology <LINK2>
F/21-56/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/61-139/Disulfide bonds: #status predicted

```

```

Query Match      7.1%; Score 103.5; DB 1; Length 354;
Best Local Similarity 30.1%; Pred. No. 0.58;
Matches 34; Conservative 21; Mismatches 37; Indels 21; Gaps 7;

Qy      31 GKNAYLPQSYTLPTSGTLVPMCWGKGF-----CPMSQCTNELRTDERNVT--YOKSS--- 81
Db      54 GGNVTLPPKFLRD-----PTAFSGTTHKIRIKWKTLPSDYLKEVDVFSWGMGHHKTYGG 107
Qy      82 ---RYQLGDNKGVSLIKNVTLDDHGYTCRIQFPGMLNDKLEKLDIK 131
Db      108 YQGRVFLKG--SDNDASLVITDITDDYGRYKCEV-IGLEDDTV-VALDIQ 157

```

Search completed: November 22, 2003, 05:44:23  
Job time : 27.1753 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 01:50:35 ; Search time 13.5189 Seconds  
(without alignments)  
977.484 Million cell updates/sec

Title: US-10-004-633-10

Perfect score: 1468  
Sequence: 1 MFSGLTNVCVLLQLLLAR.....VVEVENSNEYCYVNSQOPS 281

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	10.7	319	A33_HUMAN	Q99795 homo sapien
2	147.5	10.0	365	CKAR_HUMAN	P78310 homo sapien
3	134	9.1	398	SRB1_HUMAN	O00241 homo sapien
4	131	8.9	503	SHS1_HUMAN	P73324 h protein-t
5	129	8.8	387	SRB2_HUMAN	O991w8 homo sapien
6	128.5	8.8	499	SLB8_HUMAN	Q9h2z4 homo sapien
7	126.5	8.6	365	CKAR_MOUSE	P97792 mus musculu
8	124	8.4	306	CD80_MOUSE	Q00609 mus musculu
9	120.5	8.2	403	CD33_MOUSE	Q63994 mus musculu
10	119.5	8.1	322	ICOL_MOUSE	Q9jh38 mus musculu
11	118.5	8.1	551	SLI5_HUMAN	O15389 homo sapien
12	116	7.9	459	CD4_FABIT	P46630 oryctolagus
13	114	7.8	246	MOG_MOUSE	Q61885 mus musculu
14	112.5	7.7	463	SLI9_HUMAN	Q63994 mus musculu
15	109.5	7.5	245	CD4_SAIISC	Q29037 saimiri sci
16	108.5	7.4	245	MOG_RAT	Q63345 rattus norv
17	105	7.2	215	CTB2_RAT	P54900 rattus norv
18	105	7.2	354	PIK_FAT	P03994 rattus norv
19	104	7.1	354	PIK_HUMAN	P10915 homo sapien
20	104	7.1	513	SHS1_MOUSE	P97797 m protein-t
21	103.5	7.1	354	PLK_HORSE	Q28381 equus cabal
22	102.5	7.0	356	PLK_MOUSE	Q99495 mus musculu
23	102.5	7.0	457	CD4_RAT	P05540 rattus norv
24	102	6.9	457	CD4_MOUSE	P06332 mus musculu
25	101.5	6.9	458	CD4_MOUSE	P01730 homo sapien
26	100.5	6.8	458	CD4_CERAE	Q08338 cercopithe
27	100.5	6.8	458	CD4_PANTR	P71604 pan troglod
28	100.5	6.8	458	CD4_PANTR	P71604 pan troglod
29	100.5	6.8	506	SHS1_BOVIN	O46631 bos taurus
30	100	6.8	364	CD33_HUMAN	P20138 homo sapien
31	100	6.8	697	SLI4_HUMAN	Q961c7 homo sapien
32	99.5	6.8	458	CD4_MACFU	P79184 macaca fusc
33	99.5	6.8	458	CD4_MACMU	P16003 macaca mula

34	99.5	6.8	458	CD4_MACNE	Q08340 macaca neme
35	99	6.7	215	CTB2_HUMAN	O60939 homo sapien
36	98.5	6.7	346	EBF1_HUMAN	P98172 homo sapien
37	98.5	6.7	354	PLK_BOVIN	P55252 bos taurus
38	98.5	6.7	354	PLK_PTG	P10859 sus scrofa
39	98	6.7	575	IRL2_HUMAN	Q9h2z9 homo sapien
40	98	6.7	597	SLI1_PANTR	Q951h0 pan troglod
41	97	6.6	686	SLI8_HUMAN	Q96r12 homo sapien
42	97	6.6	1914	KML5_HUMAN	O15746 homo sapien
43	96.5	6.6	1006	BGAL_ASPNG	P29853 aspergillus
44	96	6.5	345	EBF1_MOUSE	P52795 mus musculu
45	96	6.5	569	SLI1_MOUSE	Q920g3 mus musculu

## ALIGNMENTS

RESULT 1  
A33\_HUMAN STANDARD; PRT; 319 AA.  
AC Q99795;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Cell surface A33 antigen precursor (Glycoprotein A33).  
GN GPA33.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Colon carcinoma;  
RX MEDLINE=97165045; PubMed=9012807;  
RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,  
RA Moritz R.L., Tu G.-F., Ji H., Whitehead R.H., Groenen L.C.,  
RA Scott A.M., Ritter G., Cohen L., Old L.J., Nice E.C.,  
RA Burgess A.W.;  
RT "The human A33 antigen is a transmembrane glycoprotein and a novel  
RT member of the immunoglobulin superfamily.";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).  
RL [2]  
RN POST-TRANSLATIONAL MODIFICATIONS.  
RP MEDLINE=97396159; PubMed=9245713;  
RX "Characterization of posttranslational modifications of human A33  
RA Ritter G., Cohen L.S., Nice E.C., Catimel B., Burgess A.W.,  
RA Moritz R.L., Ji H., Heath J.K., White S.J., Old L.J.,  
RA Simpson R.J.;  
RT "Characterization of posttranslational modifications of human A33  
RT antigen, a novel palmitoylated surface glycoprotein of human  
RT gastrointestinal epithelium.";  
RL Biochem. Biophys. Res. Commun. 236:682-686(1997).  
CC - FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.  
CC - SUBCELLULAR LOCATION: Type I membrane protein.  
CC - TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL  
CC EPITHELIAL AND IN 95% OF COLON CANCERS.  
CC - PTM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 KDA OF N-LINKED  
CC CARBOHYDRATE.  
CC - PTR: PALMITOYLATED.  
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC - SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
CC -----  
CC EMBL, U79725; AAC50957.1; -;  
CC Genew, HGNC:4445; GPA33.  
CC MIM, 602171; -;  
CC GO, GO:0005888; C:proteoglycan integral to plasma membrane; TAS.

DR GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG LIKE; 2.  
 DR Immunoglobulin domain; Lipoprotein; Palmitate; Glycoprotein;  
 KW Transmembrane; Signal; Antigen.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT DOMAIN 22 319  
 FT DOMAIN 22 319  
 FT TRAMEM 22 319  
 FT DOMAIN 22 319  
 FT DOMAIN 22 319  
 FT DOMAIN 22 319  
 FT DOMAIN 22 319  
 FT DISULFID 146 222  
 FT DISULFID 146 222  
 FT CARBOHYD 112 211  
 FT CARBOHYD 112 211  
 FT CARBOHYD 200 223  
 FT CARBOHYD 200 223  
 SQ SEQUENCE 319 AA; 35632 MW; 9BFC7AFA5C2408E CRC64;  
 Beery Match 10.7%; Score 157; DB 1; Length 319;  
 Beery Local Similarity 22.7%; Pred. No. 5.8e-06;  
 Matches 73; Conservative 40; Mismatches 116; Indels 92; Gaps 12;  
 QY 6 TLNCLVLLQLILANSLDEGVKNAVLPSCSYTLPTSGTLVPMCGKGFPMSCQTN 65  
 DB 11 TLCAVAVTDAISVERPPDVLASQSKSTLPCTHTSTSSR-----EGLIQNDKL-- 61  
 QY 66 ELRLDERNNVYQSKSRVQLKEDLNK-----GVSLLIKNVTLDDHGYCCRIQ 114  
 DB 62 -LTLTERVAVIWPFSKNVYHGLYKNRYSINNAEQSDASITIDLTWADNGYECV- 119  
 QY 115 PFGILNND-----KKLEKLDIKAVKTP----- 137  
 DB 120 --SLMSDEGNTKSRVRLVLPSPKPCGIGETITGNITQTLTCKSKGSPFOYSWKR 177  
 QY 138 -----AQTAGDSTTASPRTLTERNGSEFTGLVTLNHNNGTKISTWADEIKDSC 187  
 DB 178 VMLNDEQPLAPASQ--PVSLKNISTDTSG--YITTSNEBTOCINTAVARSS 232  
 QY 188 ETRIRAHIGVGSAGLTALIIIVILKWSG-----KKKLSLSLTLANLPFG 239  
 DB 233 NNVAVYGVGVVAVL---IIIGITII---YCCCRGKDDNTEDKEDARPNREAYEPE 286  
 QY 240 GLANNAVIRSEENITYTEE 260  
 DB 287 QIRELS--REREEDDYROE 305  
 RESULT 2  
 CXAR\_HUMAN STANDARD; PRT; 365 AA.  
 ID CXAR\_HUMAN STANDARD; PRT; 365 AA.  
 AC P78310; O00694;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Coxsackievirus and adenovirus receptor (Coxsackievirus B-  
 adenovirus receptor) (hCAR) (CVB3 binding protein).  
 GN CXAR OR CAR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97190109; PubMed=9036860;  
 RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,  
 RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;

RT "Isolation of a common receptor for Coxsackie B viruses and  
 RT adenoviruses 2 and 5.";  
 RT Science 275:1320-1323 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97250541; PubMed=9096397;  
 RA Tomko R.P., Xu R., Philipson L.;  
 RA "hCAR and hCAR: the human and mouse cellular receptors for subgroup C  
 RT adenoviruses and group B coxsackieviruses.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20008750; PubMed=10543405;  
 RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,  
 RA "Genomic organization and chromosomal localization of the human  
 RT Coxsackievirus B-adenovirus receptor gene.";  
 RT Hum. Genet. 105:354-359 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Anderson C.W., Kieleczawa J., Dunn J.J., Freimuth P.;  
 RT "Sequence and expression of CXAR, the human gene for the  
 RT coxsackievirus and adenovirus receptor.";  
 RT submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Anderson B., Tomko R., Anderson K., Darban H., Oncu D., Mizra M.,  
 RA Sollerbrant K., Sonhammer E., Philipson L.;  
 RT "Putative regulatory domains in the human and mouse CAR genes.";  
 RT submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cervix;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carlini P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Wuzny K.C., Sodergren E.V., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND  
 CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC or send an email to license@isb.ch).  
 CC -----  
 CC EMBL: Y07593; CAA68868.1; -  
 CC EMBL: U90716; AAC51234.1; -  
 CC EMBL: AF169366; AAF05908.1; -  
 CC EMBL: AF169360; AAF05908.1; JOINED.  
 CC EMBL: AF169361; AAF05908.1; JOINED.  
 CC EMBL: AF169362; AAF05908.1; JOINED.

DR EMBL: AF169363; AAF05908.1; JOINED.  
 DR EMBL: AF169364; AAF05908.1; JOINED.  
 DR EMBL: AF169365; AAF05908.1; JOINED.  
 DR EMBL: AF200465; AAF24344.1; -  
 DR EMBL: AF242865; AAG01088.1; -  
 DR EMBL: AF242862; AAG01088.1; JOINED.  
 DR EMBL: AF242864; AAG01088.1; JOINED.  
 DR EMBL: BC003684; AAH03684.1; -  
 DR EMBL: BC010536; AAH10536.1; -  
 DR PDB: 1FAJ, 13-JUL-01.  
 DR PDB: 1FSW, 08-NOV-00.  
 DR PDB: 1KAC, 24-NOV-99.  
 DR Genew; HGNC:2559; CXADR.  
 DR MIM; 602621; -  
 DR GO; GO:0005887; C:integral to plasma membrane; TMS.  
 DR GO; GO:0004872; F:receptor activity; TMS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR PROSITE; PSS00835; IG\_LIKE; 2.  
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;  
 KW Repeat; 3d-structure.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.  
 FT DOMAIN 20 237 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 238 258 POTENTIAL.  
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 134 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 141 128 IG-LIKE C2-TYPE 2.  
 FT DISULFID 41 120 BY SIMILARITY.  
 FT DISULFID 162 212 BY SIMILARITY.  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 365 AA; 40029 MW; AB01C6346CB7EE64 CRC64;  
 Query Match 10.0%; Score 147.5; DB 1; Length 365;  
 Best Local Similarity 23.5%; Pred. No. 4.3e-05;  
 Matches 63; Conservative 47; Mismatches 101; Indels 57; Gaps 11;  
 QY 7 LNCVLLLLQLL-LARSL-----EDGYKVEGKNAALPSCSYTLPTSGTLVPMWCKGFCPW 60  
 DB 4 LILCFVLLCGVVDPAFARSITTPPEMIEKAKGETAYLPCFKFLTS-----PEDQGPLDIEM 57  
 QY 61 SGCNELRLRDERAVTYQKSSRY-----QLKG-----DLNKGDVSLIINVTLDHGT 108  
 DB 58 LISPADQKQDVIIILYSGDKTYDDYYPDLKGRVHFTSNLKSQDASINVTNLQSLDIGN 117  
 QY 109 YCCRI-QFPGIMNDKXELKDKIAA-----KVTPAQTA-----HGDS 145  
 DB 118 YCCVKKAPGVAN-KKHLVVLVPSGARCYVDSSSEITGSPFKICEPKESLPLQYEMQ 176  
 QY 146 TTASPRTLTERRNGSETQTLVTLHNNGTKSTWADKISGSETIRTAHIGV---GVSA 202  
 DB 177 KLSQSQKMPSTWLAEMTSVSVKNASSEYSGTSCYRANRVSDDCLRLNVVPPSKA 236  
 QY 203 GLTALLIIVGI-----LKKYSCKKK 224  
 DB 237 GLIAGAIITLALALGLIIFCCRRK 264  
 RESULT 3  
 SRBI HUMAN STANDARD; PRT; 398 AA.  
 AC 000241; O8TB12; Q9H1U5; Q9Y4V0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Signal-regulatory protein beta-1 precursor (SRP-beta-1).  
 GN SRPB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97215901; PubMed=9062191;  
 RA Khaitonenkov A., Chen Z., Sures I., Wang H., Schilling J.,  
 RA Ullrich A.;  
 RT "A family of proteins that inhibit signalling through tyrosine kinase  
 RT receptors";  
 RL Nature 386:181-186(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leveasialho M.H., Leverkus M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McMahon L.J., McIay K., McKerray A.A.,  
 RA Malne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillips B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W., Thorpe A.,  
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP INTERACTION WITH TYROBP AND SYK.  
 RX PubMed=11169422;  
 RA Tomasello E., Cant C., Buehring H.-J., Vely F., Andre P., Seifert M.,  
 RA Ullrich A., Viver E.;  
 RT "Association of signal-regulatory proteins beta with KAPAP/DAP-12";  
 RL Eur. J. Immunol. 30:2147-2156(2000).  
 RN [5]







RP ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107; GLY-109 AND  
 RP VAL-111.  
 RC TISSUE=Brain;  
 RX MEDLINE=20053880; PubMed=10585853;  
 RA Sano S.-I., Ohnishi H., Kubota M.;  
 RT "Gene structure of mouse BIR/SHPS-1.";  
 RL Biochem. J. 344:667-675 (1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stebbins G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
 RA Bailey J., Barclay K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp W., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,  
 RA Coulson A., Coville G.J., Deacon R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grainger D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharalaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McComachie L.J., McIlroy J.C., Murray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillips B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,  
 RA Rice C.M., Ross M.T., Scott C.B., Senra H.R., Showkhen R., Sims S.,  
 RA Skuse C.D., Smith M.L., Soderlund C., Stewart C.A., Sultson J.E.,  
 RA Swan R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871 (2001).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS LEU-44; THR-50;  
 RP THR-51; ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107;  
 RP GLY-109 AND VAL-111.  
 RC TISSUE=Brain, Kidney, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shedenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [7]  
 RP FUNCTION AND INTERACTIONS WITH FYB, SCAP2 AND PTK2B.  
 RX MEDLINE=99401000; PubMed=10465959;  
 RA Tjime J.F., Swanson K.D., Marie-Cardine A., Raab M., Rudd C.E.,  
 RA Schraven B., Neel B.G.;  
 RT "SHPS-1 is a scaffold for assembling distinct adhesion-regulated  
 RT multi-protein complexes in macrophages.";  
 RL Curr. Biol. 9:927-930 (1999).  
 RN [8]  
 RP PHOSPHORYLATION BY JAK2, AND INTERACTIONS WITH PTEN1 AND JAK2.

RX MEDLINE=20428742; PubMed=10842184;  
 RA Stofega M.R., Argersinger L.S., Wang H., Ulrich A., Carter-Su C.;  
 RT "Negative regulation of growth hormone receptor/JAK2 signaling by  
 RT signal regulatory protein alpha.";  
 RL J. Biol. Chem. 275:28222-28229 (2000).  
 RN [9]  
 RP FUNCTION AND INTERACTION WITH CD47.  
 RX MEDLINE=21400825; PubMed=11509594;  
 RA Latour S., Tanaka H., Demure C., Mateo V., Rubio M., Brown E.J.,  
 RA Maliszewski G., Lindberg F.P., Oldenborg A., Ulrich A.;  
 RA Delapessie G., Sarrail M.;  
 RT "Bidirectional negative regulation of human T and dendritic cells by  
 RT CD47 and its cognate receptor signal-regulator protein-alpha:  
 RT down-regulation of IL-12 responsiveness and inhibition of dendritic  
 cell activation.";  
 RL J. Immunol. 167:2547-2554 (2001).  
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts  
 CC as docking protein and induces translocation of PTEN, PTEN1 and  
 CC other binding partners from the cytosol to the plasma membrane.  
 CC Supports adhesion of cerebellar neurons, neurite outgrowth and  
 CC glial cell attachment. May play a key role in intracellular  
 CC signaling during synaptogenesis and in synaptic function (by  
 CC simulating). Involved in the negative regulation of receptor  
 CC tyrosine kinase-coupled cellular responses induced by cell  
 CC adhesion, growth factors or insulin. Mediates negative regulation  
 CC of phagocytosis, mast cell activation and dendritic cell  
 CC activation. CD47 binding prevents maturation of immature dendritic  
 CC cells and inhibits cytokine production by mature dendritic cells.  
 CC -1- SUBUNIT: Binds PTEN1 when tyrosine-phosphorylated, except in  
 CC macrophages, where it primarily binds PTEN6. Binds GRB2 in vitro.  
 CC Binds FGR (by similarity). Binds JAK2 irrespective of its  
 CC phosphorylation status and forms a stable complex. Binds SCAP1  
 CC and/or SCAP2. The resulting complex recruits FYB. Binds PTK2B.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=P78324-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P78324-2; Sequence=VSP\_007030;  
 CC Note=No experimental confirmation available;  
 CC Name=3;  
 CC IsoId=P78324-3; Sequence=VSP\_007029;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain.  
 CC Detected on myeloid cells, but not T cells. Detected at lower  
 CC levels in heart, placenta, lung, testis, ovary, colon, liver,  
 CC small intestine, prostate, spleen, kidney, skeletal muscle and  
 CC pancreas.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: Phosphorylated on tyrosine residues in response to  
 CC stimulation with BGF, growth hormone, insulin and PDGF.  
 CC Dephosphorylated by PTEN1.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; D86043; BAA12974.1; -;  
 CC EMBL; Y10375; CAA71403.1; -;  
 CC EMBL; Y11047; CAA71944.1; ALT INIT.  
 CC EMBL; AB023430; BAA87929.1; -;  
 CC EMBL; AC004832; ARI19260.1; -;  
 CC EMBL; AL034562; CAB38874.1; -;  
 CC EMBL; AL049634; CAB46662.1; ALT-SEQ.  
 CC EMBL; AL117335; CAC12723.1; -;  
 CC EMBL; BC026692; AAB26692.1; -;

```

DR EMBL: BC033092; AAH33092.1; -
DR EMBL: BC038510; AAH38510.1; -
DR PIR: JCS287; JCS287.
DR Genew: HGNC:9662; PTPNS1.
DR MIM: 602461; -
DR GO: GO:0005886; Cytoplasm membrane; TAS.

Query Match      8.9%; Score 131; DB 1; Length 501;
Best Local Similarity 22.5%; Pred. No. 0.0016;
Matches 67; Conservative 45; Mismatches 124; Indels 62; Gaps 13;

QY 13 LLOLLAASL-----EDGKVEGVGNAYLPCSYT--LPTSGTLVPMCWG 54
DB 14 LLLCLLLAASCAAGVAGEELVIOIPDKSVSAAGESALHCTVSLIVG-----PIQMF 69
QY 55 KQPCFWSQCTNELLRDE-----RNVYQKSSRYQLKGLNKGDVSLIKVTLDDHGT 109
DB 70 RGAGP---ARELIYQKXGHPFRVTVVSESTRK-----NMDFSISINITPADAGTY 119
QY 110 CCRIOFPGMLNDKLE---LKLDIKAAYTPAQTHGSDTASPR---TLTTERNGSET 162
DB 120 YC-VKFRKGSPTPEFGSGAGTELISVAKPSAP--VVGGAARATPQHTVSFTCESHGFSP 176
QY 163 QTLVTLHNNNGTSTKISTWDEIKDGETRTAIHIGVAGLTLILIGVLLIKWYSCCK 222
DB 177 RQITLKFNGNELSDFOINVPDVGESVSYSHSTAKV-----VLTRDVTAS 223
QY 223 KLLSLSLITLANLPFGGLAAGAVRISSENIYTIENVEVENSENYCYVNSQOP 280
DB 224 QVICVAHVTLQGDPLRGTAANLSET-IRVPPTLEVTVQGV-BAENQVNVTCVRKPYF 279

RESULT 5
SRB2_HUMAN          STANDARD;          PRT;          387 AA.
AC Q9P1W8; Q8WMA5; Q9NOK8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Signal-regulatory protein beta-2 precursor (SIRP-beta-2) (SIRP-b2).
GN SIRPB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RX 1)
RX SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Placenta;
RX MEDLINE=21036165; PubMed=1185750;
RA Ichigotani Y., Matsuda S., Machida K., Oshima K., Iwamoto T.,
RA Yamaki K., Hayakawa T., Hamaguchi M.;
RT "Molecular cloning of a novel human gene (SIRP-B2) which encodes a new
RT member of the SIRP/SHPS-1 protein family.";
RT J. Hum. Genet. 45:378-382(2000).
[2]
RX SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Ellington A.G., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenslaio M.H., Levenslaio M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

```

```

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sultson J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thoype A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilting L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RT Nature 414:865-871(2001).
[3]
RX SEQUENCE FROM N.A. (ISOFORM 3).
RX TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probable immunoglobulin-like cell surface receptor.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9P1W8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9P1W8-2; Sequence=VSP_007027;
CC Note=No experimental confirmation available;
CC Name=3;
CC Note=No experimental confirmation available;
CC IsoId=Q9P1W8-3; Sequence=VSP_007028;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Detected in liver, and at very low levels in
CC brain, heart, lung, pancreas, kidney, placenta and skeletal
CC muscle.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C1-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
DR EMBL: AB042624; BAA95692.1; -
DR EMBL: AB138804; CAC00474.1; -
DR EMBL: BC020629; AAH20629.1; ALT_INIT.
DR Genew: HGNC:15757; SIRPB2.
DR MIM: 605466; -
DR GO: GO:0007267; P:cell-cell signaling; TAS.
DR GO: GO:0007242; P:intracellular signaling cascade; TAS.
DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.
DR InterPro: IPR007110; I9-1like.
DR InterPro: IPR003597; I9_c1.
DR InterPro: IPR003006; I9_MHC.
DR Pfam: PF00047; I9_3.

```

DR SMART; SM00407; IGc1; 2.  
DR SMART; SM00406; IGv; 1.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
KM Repeat; signal; Transmembrane; Immunoglobulin domain; Glycoprotein;  
KM Alternative splicing.  
FT SIGNAL 1 28  
FT CHAIN 29 387  
FT DOMAIN 29 360  
FT TRANSMEM 361 383  
FT DOMAIN 384 387  
FT DOMAIN 29 137  
FT DOMAIN 146 245  
FT DOMAIN 252 340  
FT DISULFID 53 119  
FT DISULFID 168 226  
FT CARBOHYD 271 329  
FT CARBOHYD 243 243  
FT CARBOHYD 268 268  
FT CARBOHYD 309 309  
FT CARBOHYD 317 317  
FT VARSPIC 1 33  
FT VARSPIC 144 360  
FT CONFLICT 263 263  
FT CONFLICT 286 286  
SQ SEQUENCE 387 AA; 42495 MW; F7F20C9F6BDE64B CMC64;  
Query Match 8.8%; Score 129; DB 1; Length 387;  
Best local similarity 23.1%; Pred. No. 0.0017;  
Matches 65; Conservative 38; Mismatches 90; Indels 88; Gaps 13;  
QY 12 LLLQLLARSLEDEYK-----VEYKNAVLPCEST--LPTSGTIVPMCKGKF 57  
DB 15 LLLTLGLTEVAGEEELQMIQPEKLLLVYGTATLACTYTSLLPVG---PYLMFRGV 70  
QY 58 CPMSQCTNELRTDERNTVYOKS---RYOLKSDL---NKGVSLIKNTVLTLDHGYTC 110  
DB 71 GP-----GRELIYNQKXGHPRTVTSVDTLRNMNDSIRISITPADVGYTY 118  
QY 111 C-----RIQF--PGLMNDKLEIKDIDK-----AAKVTPAQTAGDSTTAS 149  
DB 119 CVKFKKSPENVEFKSGPG-----TEWALGAKPSAPVGLPAARTPEHTV----- 164  
QY 150 PRITLTERNGSEQTQVTLVTHNNGTKISTADEIKDSEITRTAIHIGVGSAGITLALI 209  
DB 165 --SFTCSHGFSPPDITLKWFKNNGNELSDFTQVNDPTQSVAYSIRSTARV----- 213  
QY 210 IGVLLIKWYSCKKKSLSLITLANLPGGILAN-AGAVRI 249  
DB 214 ----VLDPMDVRSOVICVAVHTLQGDPLRGTANLSEAIRV 250  
RESULT 6  
SILE\_HUMAN  
ID SILE\_HUMAN STANDARD; PRT; 499 AA.  
AC Q9NYZ4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Sialic acid binding Ig-like lectin 8 precursor (Siglec-8)  
DE (Sialoadhesin family member-2) (SAF-2).  
GN SIGLEC8 OR SAF2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=20549027; PubMed=11095983;  
RA Fousheas G., Yousef G.M., Diamandis E.P.;  
RT "Molecular characterization of a siglec8 variant containing

RT Cytoplasmic tyrosine-based motifs, and mapping of the siglec8 gene.";  
RL Biochem. Biophys. Res. Commun. 278:775-781(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RX MEDLINE=20314554; PubMed=10856141;  
RA Kikly K.K., Bochner B.S., Freeman S.D., Tan K.B., Gallagher K.T.,  
RA D'Alessio K.J., Holmes S.D., Abrahamson J.A., Erickson-Willer C.L.,  
RA Murdoch P.R., Tachimoto H., Schleimer R.P., White J.R.;  
RT "Identification of SAF-2, a novel siglec expressed on eosinophils,  
RT mast cells, and basophils.";  
RL J. Allergy Clin. Immunol. 105:1093-1100(2000).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RX TISSUE=Bosinophil;  
RC MEDLINE=20092847; PubMed=10625619;  
RA Floyd H., Ni J., Cornish A.L., Zeng Z., Liu D., Carter K.C., Steel J.,  
RA Crocker P.R.;  
RT "Siglec-8, A novel eosinophil-specific member of the immunoglobulin  
RT superfamily.";  
RL J. Biol. Chem. 275:861-866(2000).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Floyd H., Zhang J.Q., Crocker P.R.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Putative adhesion molecule that mediates sialic-acid  
CC dependent binding to cells. Preferentially binds to alpha2,3-  
CC linked sialic acid. Also binds to alpha2,6-linked sialic acid. The  
CC sialic acid recognition site may be masked by cis interactions  
CC with sialic acids on the same cell surface.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Name=1; Synonyms=long;  
CC Name=2;  
CC Name=3;  
CC Name=4;  
CC Name=5;  
CC Name=6;  
CC Name=7;  
CC Name=8;  
CC Name=9;  
CC Name=10;  
CC Name=11;  
CC Name=12;  
CC Name=13;  
CC Name=14;  
CC Name=15;  
CC Name=16;  
CC Name=17;  
CC Name=18;  
CC Name=19;  
CC Name=20;  
CC Name=21;  
CC Name=22;  
CC Name=23;  
CC Name=24;  
CC Name=25;  
CC Name=26;  
CC Name=27;  
CC Name=28;  
CC Name=29;  
CC Name=30;  
CC Name=31;  
CC Name=32;  
CC Name=33;  
CC Name=34;  
CC Name=35;  
CC Name=36;  
CC Name=37;  
CC Name=38;  
CC Name=39;  
CC Name=40;  
CC Name=41;  
CC Name=42;  
CC Name=43;  
CC Name=44;  
CC Name=45;  
CC Name=46;  
CC Name=47;  
CC Name=48;  
CC Name=49;  
CC Name=50;  
CC Name=51;  
CC Name=52;  
CC Name=53;  
CC Name=54;  
CC Name=55;  
CC Name=56;  
CC Name=57;  
CC Name=58;  
CC Name=59;  
CC Name=60;  
CC Name=61;  
CC Name=62;  
CC Name=63;  
CC Name=64;  
CC Name=65;  
CC Name=66;  
CC Name=67;  
CC Name=68;  
CC Name=69;  
CC Name=70;  
CC Name=71;  
CC Name=72;  
CC Name=73;  
CC Name=74;  
CC Name=75;  
CC Name=76;  
CC Name=77;  
CC Name=78;  
CC Name=79;  
CC Name=80;  
CC Name=81;  
CC Name=82;  
CC Name=83;  
CC Name=84;  
CC Name=85;  
CC Name=86;  
CC Name=87;  
CC Name=88;  
CC Name=89;  
CC Name=90;  
CC Name=91;  
CC Name=92;  
CC Name=93;  
CC Name=94;  
CC Name=95;  
CC Name=96;  
CC Name=97;  
CC Name=98;  
CC Name=99;  
CC Name=100;  
CC Name=101;  
CC Name=102;  
CC Name=103;  
CC Name=104;  
CC Name=105;  
CC Name=106;  
CC Name=107;  
CC Name=108;  
CC Name=109;  
CC Name=110;  
CC Name=111;  
CC Name=112;  
CC Name=113;  
CC Name=114;  
CC Name=115;  
CC Name=116;  
CC Name=117;  
CC Name=118;  
CC Name=119;  
CC Name=120;  
CC Name=121;  
CC Name=122;  
CC Name=123;  
CC Name=124;  
CC Name=125;  
CC Name=126;  
CC Name=127;  
CC Name=128;  
CC Name=129;  
CC Name=130;  
CC Name=131;  
CC Name=132;  
CC Name=133;  
CC Name=134;  
CC Name=135;  
CC Name=136;  
CC Name=137;  
CC Name=138;  
CC Name=139;  
CC Name=140;  
CC Name=141;  
CC Name=142;  
CC Name=143;  
CC Name=144;  
CC Name=145;  
CC Name=146;  
CC Name=147;  
CC Name=148;  
CC Name=149;  
CC Name=150;  
CC Name=151;  
CC Name=152;  
CC Name=153;  
CC Name=154;  
CC Name=155;  
CC Name=156;  
CC Name=157;  
CC Name=158;  
CC Name=159;  
CC Name=160;  
CC Name=161;  
CC Name=162;  
CC Name=163;  
CC Name=164;  
CC Name=165;  
CC Name=166;  
CC Name=167;  
CC Name=168;  
CC Name=169;  
CC Name=170;  
CC Name=171;  
CC Name=172;  
CC Name=173;  
CC Name=174;  
CC Name=175;  
CC Name=176;  
CC Name=177;  
CC Name=178;  
CC Name=179;  
CC Name=180;  
CC Name=181;  
CC Name=182;  
CC Name=183;  
CC Name=184;  
CC Name=185;  
CC Name=186;  
CC Name=187;  
CC Name=188;  
CC Name=189;  
CC Name=190;  
CC Name=191;  
CC Name=192;  
CC Name=193;  
CC Name=194;  
CC Name=195;  
CC Name=196;  
CC Name=197;  
CC Name=198;  
CC Name=199;  
CC Name=200;  
CC Name=201;  
CC Name=202;  
CC Name=203;  
CC Name=204;  
CC Name=205;  
CC Name=206;  
CC Name=207;  
CC Name=208;  
CC Name=209;  
CC Name=210;  
CC Name=211;  
CC Name=212;  
CC Name=213;  
CC Name=214;  
CC Name=215;  
CC Name=216;  
CC Name=217;  
CC Name=218;  
CC Name=219;  
CC Name=220;  
CC Name=221;  
CC Name=222;  
CC Name=223;  
CC Name=224;  
CC Name=225;  
CC Name=226;  
CC Name=227;  
CC Name=228;  
CC Name=229;  
CC Name=230;  
CC Name=231;  
CC Name=232;  
CC Name=233;  
CC Name=234;  
CC Name=235;  
CC Name=236;  
CC Name=237;  
CC Name=238;  
CC Name=239;  
CC Name=240;  
CC Name=241;  
CC Name=242;  
CC Name=243;  
CC Name=244;  
CC Name=245;  
CC Name=246;  
CC Name=247;  
CC Name=248;  
CC Name=249;  
CC Name=250;  
CC Name=251;  
CC Name=252;  
CC Name=253;  
CC Name=254;  
CC Name=255;  
CC Name=256;  
CC Name=257;  
CC Name=258;  
CC Name=259;  
CC Name=260;  
CC Name=261;  
CC Name=262;  
CC Name=263;  
CC Name=264;  
CC Name=265;  
CC Name=266;  
CC Name=267;  
CC Name=268;  
CC Name=269;  
CC Name=270;  
CC Name=271;  
CC Name=272;  
CC Name=273;  
CC Name=274;  
CC Name=275;  
CC Name=276;  
CC Name=277;  
CC Name=278;  
CC Name=279;  
CC Name=280;  
CC Name=281;  
CC Name=282;  
CC Name=283;  
CC Name=284;  
CC Name=285;  
CC Name=286;  
CC Name=287;  
CC Name=288;  
CC Name=289;  
CC Name=290;  
CC Name=291;  
CC Name=292;  
CC Name=293;  
CC Name=294;  
CC Name=295;  
CC Name=296;  
CC Name=297;  
CC Name=298;  
CC Name=299;  
CC Name=300;  
CC Name=301;  
CC Name=302;  
CC Name=303;  
CC Name=304;  
CC Name=305;  
CC Name=306;  
CC Name=307;  
CC Name=308;  
CC Name=309;  
CC Name=310;  
CC Name=311;  
CC Name=312;  
CC Name=313;  
CC Name=314;  
CC Name=315;  
CC Name=316;  
CC Name=317;  
CC Name=318;  
CC Name=319;  
CC Name=320;  
CC Name=321;  
CC Name=322;  
CC Name=323;  
CC Name=324;  
CC Name=325;  
CC Name=326;  
CC Name=327;  
CC Name=328;  
CC Name=329;  
CC Name=330;  
CC Name=331;  
CC Name=332;  
CC Name=333;  
CC Name=334;  
CC Name=335;  
CC Name=336;  
CC Name=337;  
CC Name=338;  
CC Name=339;  
CC Name=340;  
CC Name=341;  
CC Name=342;  
CC Name=343;  
CC Name=344;  
CC Name=345;  
CC Name=346;  
CC Name=347;  
CC Name=348;  
CC Name=349;  
CC Name=350;  
CC Name=351;  
CC Name=352;  
CC Name=353;  
CC Name=354;  
CC Name=355;  
CC Name=356;  
CC Name=357;  
CC Name=358;  
CC Name=359;  
CC Name=360;  
CC Name=361;  
CC Name=362;  
CC Name=363;  
CC Name=364;  
CC Name=365;  
CC Name=366;  
CC Name=367;  
CC Name=368;  
CC Name=369;  
CC Name=370;  
CC Name=371;  
CC Name=372;  
CC Name=373;  
CC Name=374;  
CC Name=375;  
CC Name=376;  
CC Name=377;  
CC Name=378;  
CC Name=379;  
CC Name=380;  
CC Name=381;  
CC Name=382;  
CC Name=383;  
CC Name=384;  
CC Name=385;  
CC Name=386;  
CC Name=387;  
CC Name=388;  
CC Name=389;  
CC Name=390;  
CC Name=391;  
CC Name=392;  
CC Name=393;  
CC Name=394;  
CC Name=395;  
CC Name=396;  
CC Name=397;  
CC Name=398;  
CC Name=399;  
CC Name=400;  
CC Name=401;  
CC Name=402;  
CC Name=403;  
CC Name=404;  
CC Name=405;  
CC Name=406;  
CC Name=407;  
CC Name=408;  
CC Name=409;  
CC Name=410;  
CC Name=411;  
CC Name=412;  
CC Name=413;  
CC Name=414;  
CC Name=415;  
CC Name=416;  
CC Name=417;  
CC Name=418;  
CC Name=419;  
CC Name=420;  
CC Name=421;  
CC Name=422;  
CC Name=423;  
CC Name=424;  
CC Name=425;  
CC Name=426;  
CC Name=427;  
CC Name=428;  
CC Name=429;  
CC Name=430;  
CC Name=431;  
CC Name=432;  
CC Name=433;  
CC Name=434;  
CC Name=435;  
CC Name=436;  
CC Name=437;  
CC Name=438;  
CC Name=439;  
CC Name=440;  
CC Name=441;  
CC Name=442;  
CC Name=443;  
CC Name=444;  
CC Name=445;  
CC Name=446;  
CC Name=447;  
CC Name=448;  
CC Name=449;  
CC Name=450;  
CC Name=451;  
CC Name=452;  
CC Name=453;  
CC Name=454;  
CC Name=455;  
CC Name=456;  
CC Name=457;  
CC Name=458;  
CC Name=459;  
CC Name=460;  
CC Name=461;  
CC Name=462;  
CC Name=463;  
CC Name=464;  
CC Name=465;  
CC Name=466;  
CC Name=467;  
CC Name=468;  
CC Name=469;  
CC Name=470;  
CC Name=471;  
CC Name=472;  
CC Name=473;  
CC Name=474;  
CC Name=475;  
CC Name=476;  
CC Name=477;  
CC Name=478;  
CC Name=479;  
CC Name=480;  
CC Name=481;  
CC Name=482;  
CC Name=483;  
CC Name=484;  
CC Name=485;  
CC Name=486;  
CC Name=487;  
CC Name=488;  
CC Name=489;  
CC Name=490;  
CC Name=491;  
CC Name=492;  
CC Name=493;  
CC Name=494;  
CC Name=495;  
CC Name=496;  
CC Name=497;  
CC Name=498;  
CC Name=499;  
CC Name=500;  
CC Name=501;  
CC Name=502;  
CC Name=503;  
CC Name=504;  
CC Name=505;  
CC Name=506;  
CC Name=507;  
CC Name=508;  
CC Name=509;  
CC Name=510;  
CC Name=511;  
CC Name=512;  
CC Name=513;  
CC Name=514;  
CC Name=515;  
CC Name=516;  
CC Name=517;  
CC Name=518;  
CC Name=519;  
CC Name=520;  
CC Name=521;  
CC Name=522;  
CC Name=523;  
CC Name=524;  
CC Name=525;  
CC Name=526;  
CC Name=527;  
CC Name=528;  
CC Name=529;  
CC Name=530;  
CC Name=531;  
CC Name=532;  
CC Name=533;  
CC Name=534;  
CC Name=535;  
CC Name=536;  
CC Name=537;  
CC Name=538;  
CC Name=539;  
CC Name=540;  
CC Name=541;  
CC Name=542;  
CC Name=543;  
CC Name=544;  
CC Name=545;  
CC Name=546;  
CC Name=547;  
CC Name=548;  
CC Name=549;  
CC Name=550;  
CC Name=551;  
CC Name=552;  
CC Name=553;  
CC Name=554;  
CC Name=555;  
CC Name=556;  
CC Name=557;  
CC Name=558;  
CC Name=559;  
CC Name=560;  
CC Name=561;  
CC Name=562;  
CC Name=563;  
CC Name=564;  
CC Name=565;  
CC Name=566;  
CC Name=567;  
CC Name=568;  
CC Name=569;  
CC Name=570;  
CC Name=571;  
CC Name=572;  
CC Name=573;  
CC Name=574;  
CC Name=575;  
CC Name=576;  
CC Name=577;  
CC Name=578;  
CC Name=579;  
CC Name=580;  
CC Name=581;  
CC Name=582;  
CC Name=583;  
CC Name=584;  
CC Name=585;  
CC Name=586;  
CC Name=587;  
CC Name=588;  
CC Name=589;  
CC Name=590;  
CC Name=591;  
CC Name=592;  
CC Name=593;  
CC Name=594;  
CC Name=595;  
CC Name=596;  
CC Name=597;  
CC Name=598;  
CC Name=599;  
CC Name=600;  
CC Name=601;  
CC Name=602;  
CC Name=603;  
CC Name=604;  
CC Name=605;  
CC Name=606;  
CC Name=607;  
CC Name=608;  
CC Name=609;  
CC Name=610;  
CC Name=611;  
CC Name=612;  
CC Name=613;  
CC Name=614;  
CC Name=615;  
CC Name=616;  
CC Name=617;  
CC Name=618;  
CC Name=619;  
CC Name=620;  
CC Name=621;  
CC Name=622;  
CC Name=623;  
CC Name=624;  
CC Name=625;  
CC Name=626;  
CC Name=627;  
CC Name=628;  
CC Name=629;  
CC Name=630;  
CC Name=631;  
CC Name=632;  
CC Name=633;  
CC Name=634;  
CC Name=635;  
CC Name=636;  
CC Name=637;  
CC Name=638;  
CC Name=639;  
CC Name=640;  
CC Name=641;  
CC Name=642;  
CC Name=643;  
CC Name=644;  
CC Name=645;  
CC Name=646;  
CC Name=647;  
CC Name=648;  
CC Name=649;  
CC Name=650;  
CC Name=651;  
CC Name=652;  
CC Name=653;  
CC Name=654;  
CC Name=655;  
CC Name=656;  
CC Name=657;  
CC Name=658;  
CC Name=659;  
CC Name=660;  
CC Name=661;  
CC Name=662;  
CC Name=663;  
CC Name=664;  
CC Name=665;  
CC Name=666;  
CC Name=667;  
CC Name=668;  
CC Name=669;  
CC Name=670;  
CC Name=671;  
CC Name=672;  
CC Name=673;  
CC Name=674;  
CC Name=675;  
CC Name=676;  
CC Name=677;  
CC Name=678;  
CC Name=679;  
CC Name=680;  
CC Name=681;  
CC Name=682;  
CC Name=683;  
CC Name=684;  
CC Name=685;  
CC Name=686;  
CC Name=687;  
CC Name=688;  
CC Name=689;  
CC Name=690;  
CC Name=691;  
CC Name=692;  
CC Name=693;  
CC Name=694;  
CC Name=695;  
CC Name=696;  
CC Name=697;  
CC Name=698;  
CC Name=699;  
CC Name=700;  
CC Name=701;  
CC Name=702;  
CC Name=703;  
CC Name=704;  
CC Name=705;  
CC Name=706;  
CC Name=707;  
CC Name=708;  
CC Name=709;  
CC Name=710;  
CC Name=711;  
CC Name=712;  
CC Name=713;  
CC Name=714;  
CC Name=715;  
CC Name=716;  
CC Name=717;  
CC Name=718;  
CC Name=719;  
CC Name=720;  
CC Name=721;  
CC Name=722;  
CC Name=723;  
CC Name=724;  
CC Name=725;  
CC Name=726;  
CC Name=727;  
CC Name=728;  
CC Name=729;  
CC Name=730;  
CC Name=731;  
CC Name=732;  
CC Name=733;  
CC Name=734;  
CC Name=735;  
CC Name=736;  
CC Name=737;  
CC Name=738;  
CC Name=739;  
CC Name=740;  
CC Name=741;  
CC Name=742;  
CC Name=743;  
CC Name=744;  
CC Name=745;  
CC Name=746;  
CC Name=747;  
CC Name=748;  
CC Name=749;  
CC Name=750;  
CC Name=751;  
CC Name=752;  
CC Name=753;  
CC Name=754;  
CC Name=755;  
CC Name=756;  
CC Name=757;  
CC Name=758;  
CC Name=759;  
CC Name=760;  
CC Name=761;  
CC Name=762;  
CC Name=763;  
CC Name=764;  
CC Name=765;  
CC Name=766;  
CC Name=767;  
CC Name=768;  
CC Name=769;  
CC Name=770;  
CC Name=771;  
CC Name=772;  
CC Name=773;  
CC Name=774;  
CC Name=775;  
CC Name=776;  
CC Name=777;  
CC Name=778;  
CC Name=779;  
CC Name=780;  
CC Name=781;  
CC Name=782;  
CC Name=783;  
CC Name=784;  
CC Name=785;  
CC Name=786;  
CC Name=787;  
CC Name=788;  
CC Name=789;  
CC Name=790;  
CC Name=791;  
CC Name=792;  
CC Name=793;  
CC Name=794;  
CC Name=795;  
CC Name=796;  
CC Name=797;  
CC Name=798;  
CC Name=799;  
CC Name=800;  
CC Name=801;  
CC Name=802;  
CC Name=803;  
CC Name=804;  
CC Name=805;  
CC Name=806;  
CC Name=807;  
CC Name=808;  
CC Name=809;  
CC Name=810;  
CC Name=811;  
CC Name=812;  
CC Name=813;  
CC Name=814;  
CC Name=815;  
CC Name=816;  
CC Name=817;  
CC Name=818;  
CC Name=819;  
CC Name=820;  
CC Name=821;  
CC Name=822;  
CC Name=823;  
CC Name=824;  
CC Name=825;  
CC Name=826;  
CC Name=827;  
CC Name=828;  
CC Name=829;  
CC Name=830;  
CC Name=831;  
CC Name=832;  
CC Name=833;  
CC Name=834;  
CC Name=835;  
CC Name=836;  
CC Name=837;  
CC Name=838;  
CC Name=839;  
CC Name=840;  
CC Name=841;  
CC Name=842;  
CC Name=843;  
CC Name=844;  
CC Name=845;  
CC Name=846;  
CC Name=847;  
CC Name=848;  
CC Name=849;  
CC Name=850;  
CC Name=851;  
CC Name=852;  
CC Name=853;  
CC Name=854;  
CC Name=855;  
CC Name=856;  
CC Name=857;  
CC Name=858;  
CC Name=859;  
CC Name=860;  
CC Name=861;  
CC Name=862;  
CC Name=863;  
CC Name=864;  
CC Name=865;  
CC Name=866;  
CC Name=867;  
CC Name=868;  
CC Name=869;  
CC Name=870;  
CC Name=871;  
CC Name=872;  
CC Name=873;  
CC Name=874;  
CC Name=875;  
CC Name=876;  
CC Name=877;  
CC Name=878;  
CC Name=879;  
CC Name=880;  
CC Name=881;  
CC Name=882;  
CC Name=883;  
CC Name=884;  
CC Name=885;  
CC Name=886;  
CC Name=887;  
CC Name=888;  
CC Name=889;  
CC Name=890;  
CC Name=891;  
CC Name=892;  
CC Name=893;  
CC Name=894;  
CC Name=895;  
CC Name=896;  
CC Name=897;  
CC Name=898;  
CC Name=899;  
CC Name=900;  
CC Name=901;  
CC Name=902;  
CC Name=903;  
CC Name=904;  
CC Name=905;  
CC Name=906;  
CC Name=907;  
CC Name=908;  
CC Name=909;  
CC Name=910;  
CC Name=911;  
CC Name=912;  
CC Name=913;  
CC Name=914;  
CC Name=915;  
CC Name=916;  
CC Name=917;  
CC Name=918;  
CC Name=919;  
CC Name=920;  
CC Name=921;  
CC Name=922;  
CC Name=923;  
CC Name=924;  
CC Name=925;  
CC Name=926;  
CC Name=927;  
CC Name=928;  
CC Name=929;  
CC Name=930;  
CC Name=931;  
CC Name=932;  
CC Name=933;  
CC Name=934;  
CC Name=935;  
CC Name=936;  
CC Name=937;  
CC Name=938;  
CC Name=939;  
CC Name=940;  
CC Name=941;  
CC Name=942;  
CC Name=943;  
CC Name=944;  
CC Name=945;  
CC Name=946;  
CC Name=947;  
CC Name=948;  
CC Name=949;  
CC Name=950;  
CC Name=951;  
CC Name=952;  
CC Name=953;  
CC Name=954;  
CC Name=955;  
CC Name=956;  
CC Name=957;  
CC Name=958;  
CC Name=959;  
CC Name=960;  
CC Name=961;  
CC Name=962;  
CC Name=963;  
CC Name=964;  
CC Name=965;  
CC Name=966;  
CC Name=967;  
CC Name=968;  
CC Name=969;  
CC Name=970;  
CC Name=971;  
CC Name=972;  
CC Name=973;  
CC Name=974;  
CC Name=975;  
CC Name=976;  
CC Name=977;  
CC Name=978;  
CC Name=979;  
CC Name=980;  
CC Name=981;  
CC Name=982;  
CC Name=983;  
CC Name=984;  
CC Name=985;  
CC Name=986;  
CC Name=987;  
CC Name=988;  
CC Name=989;  
CC Name=990;  
CC Name=991;  
CC Name=992;  
CC Name=993;  
CC Name=994;  
CC Name=995;  
CC Name=996;  
CC Name=997;  
CC Name=998;  
CC Name=999;  
CC Name=1000;  
CC Name=1001;  
CC Name=1002;  
CC Name=1003;  
CC Name=1004;  
CC Name=1005;  
CC Name=1006;  
CC Name=1007;  
CC Name=1008;  
CC Name=1009;  
CC Name=1010;  
CC Name=1011;  
CC Name=1012;  
CC Name=1013;  
CC Name=1014;  
CC Name=1015;  
CC Name=1016;  
CC Name=1017;  
CC Name=1018;  
CC Name=1019;  
CC Name=1020;  
CC Name=1021;  
CC Name=1022;  
CC Name=1023;  
CC Name=1024;  
CC Name=1025;  
CC Name=1026;  
CC Name=1027;  
CC Name=1028;  
CC Name=1029;  
CC Name=1030;  
CC Name=1031;  
CC Name=1032;  
CC Name=1033;  
CC Name=1034;  
CC Name=1035;  
CC Name=1036;  
CC Name=1037;  
CC Name=1038;  
CC Name=1039;  
CC Name=10



Db 146 CFVDSGEIGNDFKCEPKESGSLPQFEMQKLSDSQTMPTPLAEMTSPVTSVKNA5SE 205  
 Qy 157 RNSGETQTLVTHNNNGTKISTMADEIKDSEGTIRTHIHG-VGVSAGLTAL-IQVL 213  
 Db 206 YSGTYS---CTVQNRVG-----SDQCMRLDVPVPPSRAGTIAGAVITLALVIGAI 256  
 Qy 214 ILKWSCKKKK 224  
 Db 257 L---FCCHRRK 264

RESULT 8  
 ID CD80\_MOUSE STANDARD; PRT; 306 AA.  
 AC 000609;  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE T lymphocyte activation antigen CD80 precursor (activation B7-1 antigen) (B7).  
 GN CD80 OR B7.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=B-cell;  
 RX MEDLINE=91341422; PubMed=1714935;  
 RA Gray G.S., Freeman G.J., Gimml C.D., Lombard D.B., Zhou L.J., White M., Fingerhuth J.D., Nadler L.M.,  
 RT "Structure, expression, and T cell costimulatory activity of the murine homologue of the human B lymphocyte activation antigen B7.",  
 RL J. Exp. Med. 174:625-631 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=B-cell;  
 RX MEDLINE=93307789; PubMed=7686531;  
 RA Selvakumar A., White P.C., Dupont B.,  
 RT "Genomic organization of the mouse B-lymphocyte activation antigen B7.",  
 RL Immunogenetics 38:292-295 (1993).  
 CC -1- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON ACTIVATED B CELLS, GAMMA INTERFERON STIMULATED MONOCYTES AND NONCIRCULATING B-CELL MALIGNANCIES.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BETWEEN 4 AND 12 HOURS POST-ACTIVATION. PROTEIN WAS DETECTED AT CELL SURFACE AT 24 HOURS AND IT'S EXPRESSION WAS MAXIMAL FROM 48 TO 72 HOURS POST-ACTIVATION.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 CC EMBL, X60958; CAA43291.1; -;  
 DR EMBL, L12589; AAA37240.1; ALT\_SEQ.  
 DR EMBL, L12585; AAA37240.1; JOINED.  
 DR EMBL, L12586; AAA37240.1; JOINED.  
 DR EMBL, L12587; AAA37240.1; JOINED.  
 DR EMBL, L12588; AAA37240.1; JOINED.  
 DR PIR, I49503; I49503.  
 DR HSSP, P33681; IDR9.

DR MGD; MG1.101775; C880.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00409; Ig; 1.  
 DR PROSITE; PS50835; Ig\_LIKE; 2.  
 KW Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane; Receptor.  
 FT SIGNAL 1 37  
 FT CHAIN 38 306  
 FT DOMAIN 38 246  
 FT TRANSMEM 247 268  
 FT DOMAIN 269 306  
 FT DOMAIN 38 135  
 FT DOMAIN 148 229  
 FT DOMAIN 227 246  
 FT DISULFID 54 119  
 FT DISULFID 165 219  
 FT CARBOHYD 93 93  
 FT CARBOHYD 99 99  
 FT CARBOHYD 149 149  
 FT CARBOHYD 189 189  
 FT CARBOHYD 210 210  
 FT CARBOHYD 214 214  
 SQ SEQUENCE 306 AA; 34589 MW; 1DBADE0931B84C62 CRC64;

Query Match 8.44; Score 124; DB 1; Length 306;  
 Best Local Similarity 24.04; Pred. No. 0.0033;

Matches 67; Conservative 32; Mismatches 90; Indels 90; Gaps 12;

Qy 11 LLLQLLLAR-----SLGQYKVEGKNVLPCTSYTLPTSG-----TLVPM 51  
 Db 21 LILLFVLLILSLQSSVNDQSLKSVQDKLLPFRYNSPHDESEDIYQKHQKVLSV 80  
 Qy 52 CWGKGFCPSQCTNELIRTERVNTYQSSRYQLKGLNKGVSLLIKNVTLDHGYTC 111  
 Db 81 IAGK-LKVMPEYKN---RTLYDNTTY-----SLIILGLVLSRGTYS 119  
 Qy 112 RIQFP--GLMNDKLEU-KLDIRAKTTPAQTANGDSTTASPR-----IS---TWAEIK 184  
 Db 120 VQKKRGTYEVKRLALVKLSIKADFSTPNITSGNSADTKRITCFASGFPKPRFSL 179  
 Qy 152 -----TLTERNGSETQTLVTLHNNNGTK-----IS---TWAEIK 184  
 Db 180 ENGEELPQINTTISQDESELEYTSSQDPNTTRNHTIKCLIKGDAHVSDFTWKPRP 239  
 Qy 185 DSGETIRTAIHGVSAGLTALIIIGVILKWSCKKK 223  
 Db 240 DPPSKNTLVLFAGFGAVITVVVI--VVIKCF-CGR 275

RESULT 9  
 ID CD33\_MOUSE STANDARD; PRT; 403 AA.  
 AC 063964; 063997;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myeloid cell surface antigen CD33 precursor (Siglec-3).  
 GN CD33.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 33-A AND 33-B).  
 RC STRAIN=BALB/c; TISSUE=Bone marrow;  
 RX MEDLINE=94250900; PubMed=8193354;  
 RA Tchillian E.Z., Beverley P.C., Young B.D., Watt S.M.,  
 RT "Molecular cloning of two isoforms of the murine homologue of the myeloid CD33 antigen.",  
 RL Blood 83:3188-3198 (1994).

CC -!- FUNCTION: Putative adhesion molecule of myelomonocytic-derived  
 CC cells that mediates stialic-acid dependent binding to cells.  
 CC Preferentially binds to alpha2,6-linked stialic acid (by  
 CC sialinarity). The stialic acid recognition site may be masked by cis  
 CC interactions with stialic acids on the same cell surface.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=33-B;  
 CC IsoId=O63994-1; Sequence=Displayed;  
 CC Name=33-A;  
 CC IsoId=O63994-2; Sequence=VSP 002534;  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. SINGLE  
 CC (STATIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S71345; AAB30842.1; -;  
 CC EMBL; S71403; AAB30843.2; -;  
 CC MGI; MGI:99440; CD33.  
 CC InterPro; IPR007110; IG-1-like.  
 CC InterPro; IPR003599; IG.  
 CC InterPro; IPR003006; IG\_MHC.  
 CC Pfam; PF00047; Ig\_2.  
 CC SMART; SM00409; Ig\_1.  
 CC PROSITE; PS50835; IG\_LIKE; 1.  
 CC Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;  
 CC Immunoglobulin domain; Alternative splicing.  
 CC SIGNAL 1 16  
 CC CHAIN 17 403 MYELOID CELL SURFACE ANTIGEN CD33.  
 CC DOMAIN 18 240 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 241 267 POTENTIAL.  
 CC DOMAIN 268 403 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 17 120 IG-LIKE V-TYPE.  
 CC DOMAIN 145 228 IG-LIKE C2-TYPE.  
 CC DISULFID 36 169 BY SIMILARITY.  
 CC DISULFID 41 100 BY SIMILARITY.  
 CC DISULFID 163 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC VARSPLIC 287 403 ROBERTSYNHCUSPTASDAVTEGCSIHRIISRPCTALR  
 CC IOOPYRTHLRNNAVSTILRPFWISWGSLSRSTORSCTALC  
 CC SPVKNLCPIPLVPDNCICPLIPEMWMLCVSLTSLTS -> AH  
 CC OODSKVSNPNENRPLQKDSPOEOSVHTKISLDFMGGRPO  
 CC EYSEI (in isoform 33-A).  
 CC /FTId=VSP 002534.  
 CC SEQUENCE 403 AA; 44824 MW; F1FE6D5C393F0F1 CRC64;  
 CC  
 CC Query Match 8.2%; Score 120.5; DB 1; Length 403;  
 CC Best Local Similarity 23.7%; Pred. No. 0.0091;  
 CC Matches 65; Conservative 32; Mismatches 116; Indels 61; Gaps 8;

QY 155 -----TENNGSETQTLVTLHNNNGKISTWADIKDSGFTIRAHIGVSGA----- 203  
 DB 184 ITSLSRTTDSVLTFTPOPDHGKTL-TCLVTFSGAGVTVERTQLNVTKSGGMRELIV 242  
 QY 204 -----LTLLALIGVILIKWYSCCKKKLS 226  
 DB 243 LVANGAEATVKLLILGLCLVFLIWMCRKRTKLS 276  
 RESULT 10  
 ID ICOL MOUSE STANDARD; PRT; 322 AA.  
 AC 09JH8.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE ICOS ligand precursor (B7 homolog 2) (B7-H2) (B7-like protein G150)  
 DE (B7-related protein-1) (B7RP-1) (ICOS).  
 GN ICOSL OR B7H2 OR B7RP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.  
 RC TISSUE=Lymphocytes;  
 RX MEDLINE=20083495; PubMed=10617205;  
 RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmiento U., Guo J.,  
 RA Horan T., Shih G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,  
 RA Branck D., Campbell P., Chang D., Chiu L., Dai T., Duncan G.,  
 RA Elliott G.S., Hui A., McCabe S.M., Scully S., Shabinian A.,  
 RA Shaklee C.L., Van G., Mak T.W., Senaldi G.,  
 RT "T-cell co-stimulation through B7RP-1 and ICOS."  
 RL Nature 402:827-832 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Thymus;  
 RX MEDLINE=20015817; PubMed=10549624;  
 RA Swallow M.M., Wallin J.U., Sha W.C.,  
 RT "B7h, a novel costimulatory homology of B7.1 and B7.2, is induced by  
 RT TNFalpha."  
 RN Immunity 11:423-432 (1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C3H/HeJ; TISSUE=Fetal thymus;  
 RX MEDLINE=20126021; PubMed=10657606;  
 RA Ling V., Wu P.W., Fimmerty H.F., Bean K.M., Spaulding V., Fouser L.A.,  
 RA Leonard J.P., Hunter S.E., Zollner R., Thomas J.L., Miyashiro J.S.,  
 RA Jacobs K.A., Collins M.,  
 RT "Identification of G150, a novel B7-like protein that functionally  
 RT binds to ICOS receptor."  
 RL J. Immunol. 164:1653-1657 (2000).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Peritoneal blood lymphocytes;  
 RX MEDLINE=21286479; PubMed=11390480;  
 RA Ling V., Wu P.W., Miyashiro J.S., Marusic S., Fimmerty H.F.,  
 RA Collins M.,  
 RT "Differential expression of inducible costimulator-ligand splice  
 RT variants: lymphoid regulation of mouse g150-b and human g150  
 RT molecules."  
 RL J. Immunol. 166:7300-7308 (2001).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RA Ling V., Dunussi-Joannopoulos K.,  
 RT "G150 molecules and uses therefor."  
 RL Patent number WO0121796, 29-MAR-2001.  
 CC -!- FUNCTION: LIGAND FOR THE T-CELL-SPECIFIC CELL SURFACE RECEPTOR  
 CC CYTOKINE SECRETION; INDUCES ALSO B-CELL PROLIFERATION AND  
 CC DIFFERENTIATION INTO PLASMA CELLS. COULD PLAY AN IMPORTANT ROLE IN  
 CC MEDIATING LOCAL TISSUE RESPONSES TO INFLAMMATORY CONDITIONS, AS  
 CC WELL AS IN MODULATING THE SECONDARY IMMUNE RESPONSE BY CO-



CC STIMULATING MEMORY T CELL FUNCTION. DURING PREGNANCY, MAY FUNCTION  
 CC TO SKEW THE CYTOKINE OF MATERNAL T-CELLS TOWARD IMMUNOPROTECTIVE  
 CC TH2 PHENOTYPE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=Q9JH8-1; Sequence=Displayed;  
 CC Name=2; Synonyms=B;  
 CC IsoId=Q9JH8-2; Sequence=VSP\_002521;  
 CC -1- TISSUE SPECIFICITY: ISOFORM 1 HIGHEST EXPRESSION IN LYMPHOID  
 CC TISSUES, SUCH AS SPLEEN (MOSTLY IN THE MARGINAL ZONE), LYMPH NODES  
 CC (PARTICULARLY IN THE CORTX AND IN BOTH PRIMARY AND SECONDARY  
 CC FOLLICLES), THYMUS (PREDOMINANTLY IN THE MEDULLA) AND Peyer's  
 CC PATCHES (MOSTLY IN THE FOLLICLES), LOWER LEVELS IN MANY  
 CC NONLYMPHOID TISSUES, SUCH AS BRAIN, HEART, KIDNEY, LIVER, LUNG,  
 CC SKELETAL MUSCLE AND TESTIS. PRESENT ON FRESHLY ISOLATED SPLENIC B-  
 CC CELLS, T-CELLS, DENDRITIC CELLS AND MACROPHAGES. THE EXPRESSION OF  
 CC ISOFORM 2 IS RESTRICTED TO HEART, SPLEEN AND KIDNEY.  
 CC -1- DEVELOPMENTAL STAGE: DETECTED EARLY IN EMBRYOS; IN THE YOLK  
 CC SAC AT 11.5 AND 12.5 DPC AND, TO A LESSER EXTENT, IN THE LIVER AT  
 CC 14.5 DPC.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V2-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF216747; AAF45149.1; -  
 CC EMBL; AF199027; AAF34738.1; -  
 CC EMBL; AX100591; CAC36463.1; -  
 CC EMBL; AX100593; CAC36464.1; -  
 CC EMBL; AF394451; AAK77544.1; -  
 CC MGI; MGI:1354701; ICOS1.  
 CC DR GO; GO:000576; C:extracellular; IDA.  
 CC DR GO; GO:0016021; C:integral to membrane; NAS.  
 CC DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 CC DR GO; GO:0005102; F:receptor binding activity; TAS.  
 CC DR GO; GO:0045190; P:isotype switching; NAS.  
 CC DR GO; GO:0042104; P:positive regulation of activated T-cell pro. . .; TAS.  
 CC DR GO; GO:0045404; P:positive regulation of interleukin-4 biosyn. . .; IMP.  
 CC DR GO; GO:0007165; P:signal transduction; NAS.  
 CC DR GO; GO:0042110; P:T-cell activation; NAS.  
 CC DR InterPro; IPR007110; Ig-like.  
 CC DR InterPro; IPR003599; IG.  
 CC DR InterPro; IPR003006; IG\_MHC.  
 CC DR Pfam; PF00047; Ig\_1.  
 CC DR SMART; SM00409; IG\_1.  
 CC DR PROSITE; PS50835; IG\_LIKE; 2.  
 CC DR B-cell activation; Immune response; Glycoprotein;  
 CC Immunoglobulin domain; Signal; Transmembrane; Multigene family;  
 CC KM Alternative splicing.  
 CC FT SIGNAL 1 46 BY SIMILARITY.  
 CC FT CHAIN 47 322 ICOS LIGAND.  
 CC FT DOMAIN 47 277 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSMEM 278 298 POTENTIAL.  
 CC FT DOMAIN 299 322 CYTOPLASMIC (POTENTIAL).  
 CC FT DOMAIN 49 149 IG-LIKE V-TYPE.  
 CC FT DOMAIN 167 263 IG-LIKE C2-TYPE.  
 CC FT DOMAIN 31 38 POLY-LBU.  
 CC FT DOMAIN 289 292 POLY-ALA.  
 CC FT DISULFID 62 138 POTENTIAL.  
 CC FT DISULFID 185 243 POTENTIAL.  
 CC FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT

FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 321 322 HA -> TWAPVPDYLLPRYMSCLTRGLP (in  
 FT isoform 2).  
 FT /FTID=VSP\_002521.  
 FT CONFLICT 237 237 R -> H (IN REF. 4 AND 5; CAC36464).  
 FT SEQUENCE 322 AA; 35960 MW; 55CCBA4AD12E47E6 CMC64;  
 SQ  
 Query Match 8.1%; Score 119.5; DB 1; Length 322;  
 Beest Local Similarity 28.8%; Pred. No. 0.0084;  
 Matches 59; Conservative 20; Mismatches 75; Indels 51; Gaps 11;  
 QY 2 FSGLTNCVLLDLLLARSLEDGKVEGKAVLPC-----SYLFTSGTLVPMCKG 55  
 DB 28 FSGIGL--FLILLSSLCASAEYEGVAMGVSNNVLSICIDPHRRHNTL--SGLYV----- 77  
 QY 56 GFCWMOCTNELRTDERNTY-----QKSSRYLKGDLN-----KGDVSLIKNV 101  
 DB 78 ---TWQ-----TENPEVSTYTYLPYKSPGINVDSYKRGHLSDSKQGNFSLYLKOV 128  
 QY 102 TLDDHGYCCRIOPPGLMND-----KKLE--LKLDIKAKYTPAQTARHGDSTTASPRIT 154  
 DB 129 TPQDTQETCRV-----FNTATLTKLLEEVRLRVANFTVPVISTSDSSNPGQERTYT 184  
 QY 155 -TERNGSETOTLVTHNNNGTKIST 178  
 DB 185 CMSKNGYPEPVLWINTDNLIDT 209  
 RESULT 11  
 SILE\_HUMAN STANDARD; . PRT; 551 AA.  
 ID SILE\_HUMAN  
 AC A015389;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Sialic acid binding Ig-like lectin 5 precursor (Siglec-5) (obesity-  
 DE binding protein 2) (OB-BP2) (CD33 antigen-like  
 DE 2) (CD170 antigen).  
 OS SIGLECS OR OB-BP2 OR CD33L2.  
 GN Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Macrophage;  
 RX MEDLINE=98402360; PubMed=9731071;  
 RA Cornish A.L., Freeman S., Forbes G., Ni J., Zhang M.,  
 RA Gentz R., Augustus M., Carter K.C., Crocker P.R.;  
 RT "Characterization of siglec-5, a novel glycoprotein expressed on  
 RT myeloid cells related to CD33.";  
 RL Blood 92:2123-2132(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Erythrocytopenia;  
 RX MEDLINE=99357812; PubMed=10428856;  
 RA Patel N., Brinkman-Van der Linden E.C.M., Altmann S.W., Gish K.C.,  
 RA Balasubramanian S., Tiams J.C., Peterson D., Bell M.P., Bazan J.F.,  
 RA Varki A., Kastelein R.A.;  
 RT "OB-BP1/Siglec-6, A leptin- and sialic acid-binding protein of the  
 RT immunoglobulin superfamily.";  
 RL J. Biol. Chem. 274:22729-22738(1999).  
 RN [3]  
 RP ERRATUM.  
 RA Patel N., Brinkman-Van der Linden E.C.M., Altmann S.W., Gish K.C.,  
 RA Balasubramanian S., Tiams J.C., Peterson D., Bell M.P., Bazan J.F.,  
 RA Varki A., Kastelein R.A.;  
 RL J. Biol. Chem. 274:28058-28058(1999).  
 RN [4]





DR EMBL; M92840; AAA3198.1; -  
 DR PIR; A46254; A46254.  
 DR HSSP; P01730; 1WER.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding activity; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007159; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR000973; CD4-TDAG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG-MHC.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
 FT SIGNAL 1 25  
 FT CHAIN 26 459  
 FT DOMAIN 26 396  
 FT TRANSSEM 397 419  
 FT DOMAIN 420 459  
 FT DOMAIN 26 129  
 FT DOMAIN 130 208  
 FT DOMAIN 209 318  
 FT DOMAIN 319 374  
 FT CARBOHYD 299 299  
 FT DISULFID 41 113  
 FT LIPID 420 420  
 FT LIPID 423 423  
 FT SEQUENCE 459 AA; 50886 MW; B323311CBDA0013D CRC64;  
 Query Match 7.9%; Score 116; DB 1; Length 459;  
 Best Local Similarity 26.8%; Pred. No. 0.026; Indels 48; Gaps 7;  
 Matches 45; Conservative 21; Mismatches 54; Indels 48; Gaps 7;  
 QY 9 CVLLLLDQLLARSLEDGYKVEGKNAVLPSCYTLPTSGTLVPMCGKGFCEWSGQTNELL 68  
 DB 9 CLLLVLPALPLPAATWGTGTVRGK-----AGAIYELP-----CQSSQKRNVSF 51  
 QY 69 RTDERNNTY-----QKSSRYQLKGD-----LAKGVSLIKNTLDDHGT 109  
 DB 52 NWKHANOVKILGNQSSSSSFWLGNNSPLSNRVESKKNNMDGSPFLVIXDLRMDDSGT 111  
 QY 110 CCRLOPFLANDKXLEKLDIKAKVTP-AQTAGDSTTASPRITTE 156  
 DB 112 ICEV-----GDKKMEVLLVFRILTANDVTRLHQSLS-----TLTLE 148  
 RESULT 13  
 MCG\_MOUSE STANDARD; PRT; 246 AA.  
 ID MCG\_MOUSE  
 AC Q61885; P70364; O62003;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Myelin-oligodendrocyte glycoprotein precursor.  
 GN MCG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129;  
 RX MEDLINE=95130110; PubMed=7829100;  
 RA Dabdas P., Pham-Dinh D., Dautigny A.;  
 RT "Structure and polymorphism of the mouse myelin/oligodendrocyte

RT glycoprotein gene.";  
 RL Genomics 23:36-41(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gardiner M.V., Matchieu J.M.;  
 RT "Murine and human MCG are highly conserved: cDNA analysis.";  
 RL Trans. Am. Soc. Neurochem. 24:234-234(1993).  
 RN [3]  
 RP SEQUENCE OF 29-246 FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=93376728; PubMed=8367453;  
 RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,  
 RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,  
 RA Dautigny A.;  
 RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the  
 RT immunoglobulin superfamily encoded within the major  
 RT histocompatibility complex.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).  
 RN [4]  
 RP SEQUENCE OF 29-54.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=92218912; PubMed=1373175;  
 RA Amiguet P., Gardiner M.V., Zanetta J.-P., Matchieu J.-M.;  
 RT "Purification and partial structural and functional characterization  
 RT of mouse myelin/oligodendrocyte glycoprotein.";  
 RL J. Neurochem. 58:1676-1682(1992).  
 CC -1- FUNCTION: MINOR COMPONENT OF THE MYELIN SHEATH. MAY BE INVOLVED IN  
 CC COMPLETION AND/OR MAINTENANCE OF THE MYELIN SHEATH AND IN CELL-  
 CC CELL COMMUNICATION.  
 CC -1- SUBUNIT: MAY FORM HOMODIMERS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CNS. WHERE IT IS  
 CC LOCALIZED ON THE SURFACE OF MYELIN AND OLIGODENDROCYTE CYTOPLASMIC  
 CC MEMBRANES.  
 CC -1- DISEASE: REDUCED CONCENTRATIONS OF MCG ARE OBSERVED IN JIMPY AND  
 CC QUACKING DYSMYELINATING MUTANT MICE.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MCG  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- CAUTION: DO NOT CONFUSE MYELIN-OLIGODENDROCYTE GLYCOPROTEIN (MCG)  
 CC WITH OLIGODENDROCYTE-MYELIN GLYCOPROTEIN (OMG).  
 CC -----  
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L29503; AAC42023.1; -  
 DR EMBL; L29498; AAC42023.1; JOINED.  
 DR EMBL; L29500; AAC42023.1; JOINED.  
 DR EMBL; L29501; AAC42023.1; JOINED.  
 DR EMBL; L29499; AAC42023.1; JOINED.  
 DR EMBL; L29502; AAC42023.1; JOINED.  
 DR EMBL; U64572; AAB08096.1; -  
 DR EMBL; L20942; AAA03180.1; -  
 DR PIR; A55717; A55717.  
 DR MGI; MGI:97435; MCG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG-MHC.  
 DR InterPro; IPR003596; IG\_V.  
 DR Pfam; PF00047; Ig\_1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 246  
 FT DOMAIN 29 156  
 FT TRANSSEM 157 177  
 FT DOMAIN 178 209  
 FT TRANSSEM 210 230  
 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.  
 EXTRACELLULAR (POTENTIAL).  
 CYTOPLASMIC (POTENTIAL).  
 POTENTIAL.

FT DOMAIN 231 246 - EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 31 144 IG-LIKE V-TYPE.  
FT DISULFID 52 126 POTENTIAL.  
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 21 21 L -> IL (IN REF. 2).  
FT CONFLICT 32 32 R -> G (IN REF. 4).  
FT CONFLICT 95 95 G -> E (IN REF. 3).  
FT CONFLICT 169 169 P -> S (IN REF. 2).  
SQ SEQUENCE 246 AA, 28271 MM, 1F1A8A4A0D5CFE89, CRC64;  
  
Query Match 7.8%; Score 114; DB 1; Length 246;  
Best Local Similarity 27.4%; Pred. No. 0.018;  
Matches 37; Conservative 30; Mismatches 52; Indels 16; Gaps 5;  
  
QY 7 LNCVLLLDLQLLA-----RSLEDGYVE--VGNAYLPGSYTLPTISGTLVPMCKGKFCP 59  
DB 14 LSLILLLLQLSCSYACQAFRTIGYPIRALVGEALPCFISPGKNATGMEVGYRS--P 71  
QY 60 WSGCTNELLRTDERNNTYKSSRYQ-----LKGDLNKGVSLLIKNVTLDHGTTCRQ 114  
DB 72 FSRVVA--LYRNGKDDARQAPRYRGRRTLELKTETISEGVTLRITQVRFPSDSGYTCFPR 129  
QY 115 FPGLMNDKLELTLTD 129  
DB 130 DHSYGEAAMELKE 144  
  
RESULT 14  
SIL9 HUMAN STANDARD; PRT; 463 AA.  
AC 09Y336; 09BY19;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DB Static acid-binding Ig-like lectin 9 precursor (Siglec-9) (FOAP-9 protein).  
GN SIGLEC9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BONE marrow;  
RX MEDLINE=2036134; PubMed=10903842;  
RA Fousias G., Yousef G.M., Diamandis E.P.;  
RT "Identification and molecular characterization of a novel member of the siglec family (SIGLEC9).";  
RL Genomics 67:171-178(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20357333; PubMed=10801862;  
RA Zhang J.Q., Nicoll G., Jones C., Crocker P.R.;  
RT "Siglec-9, a novel static acid binding member of the immunoglobulin superfamily expressed broadly in human blood leukocytes.";  
RL J. Biol. Chem. 275:22121-22126(2000).  
RN [3]  
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-120.  
RC TISSUE=Periphereal blood;  
RX MEDLINE=20357334; PubMed=10801860;  
RA Angata T., Varki A.;  
RT "Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with static acid synthesis pathways.";  
RL J. Biol. Chem. 275:22127-22135(2000).  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANT GLU-315.  
RA Takayama K., Fujii Y., Tsuritani K., Naitou K., Kawaguchi A., Uka Y., Amemura C., Yajima Y., Yazaki M.;  
RT "Molecular cloning of a novel gene, FOAP-9, which are induced by oxylized LDL in human macrophages";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]

RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J., Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Andrease T., Trankheim M., Attix C., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;  
RL "Sequence analysis of chromosome 19q13.4.";  
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Putative adhesion molecule that mediates static-acid dependent binding to cells. Preferentially binds to alpha2,3- or 2,6-linked static acid. The static acid recognition site may be masked by cis interactions with static acids on the same cell surface.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed by peripheral blood leukocytes (neutrophils and monocytes but not eosinophils). Found in liver, fetal liver, bone marrow, placenta, spleen and in lower levels in skeletal muscle, fetal brain, stomach, lung, thymus, prostate, brain, mammary, adrenal gland, colon, trachea, cerebellum, testis, small intestine and spinal cord.  
CC -1- DOMAIN: Contains an intracytoplasmic motif referred as immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular functions as the termination of the immune response.  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. SIGLEC (STATIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AF135027; AAD26428.2; -  
CC EMBL; AF247180; AAF87223.1; -  
CC EMBL; AF227924; AAF71455.1; -  
CC EMBL; AB026265; BAB41100.1; -  
CC EMBL; AC011473; AAG23261.1; -  
CC Genew; HGNC:10878; SIGLEC9.  
CC MIM; 605640; -  
DR GO; GO:0005887; C:integral to plasma membrane; NAS.  
DR GO; GO:0005530; F:lectin; NAS.  
DR GO; GO:0007165; P:cell surface receptor linked signal transdu. .; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig\_3.  
DR SMART; SM00409; Ig; 2.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
KW Cell adhesion; Lectin; Antigen; Transmembrane; Signal; Glycoprotein; Immunoglobulin domain; Repeat; Polymorphism.  
KW SIGNAL 1 17  
FT CHAIN 18 463  
FT DOMAIN 31 348  
FT TRANSMEM 349 369  
FT DOMAIN 370 463  
FT DOMAIN 20 140  
FT DOMAIN 146 229  
FT DOMAIN 236 336  
FT SITE 431 436  
FT SITE 454 459  
FT DISULFID 36 170  
FT DISULFID 41 102  
FT DISULFID 164 213  
FT DISULFID 272 320  
FT CARBOHYD 101 101  
N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 100 100 K-> E (in dbSNP:2075803).
FT VARIANT 125 125 S-> N (in dbSNP:200658).
FT VARIANT 147 147 /FTid=VAR_014255.
FT VARIANT 147 147 N-> K (in dbSNP:273687).
FT VARIANT 315 315 /FTid=VAR_014256.
FT VARIANT 316 315 A-> E (in dbSNP:2258983).
FT VARIANT 316 316 /FTid=VAR_014257.
FT MUTAGEN 120 120 A-> D (in dbSNP:273688).
FT CONFLICT 269 269 R-> K; LOSS OF SIMILIC ACID BINDING.
SQ SEQUENCE 463 AA; 50081 MM; 807BDCED018191F2 CRC64;

```

```

Query Match 7.7%; Score 112.5; DB 1; Length 463;
Best Local Similarity 21.6%; Pred. No. 0.051;
Matches 55; Conservative 42; Mismatches 95; Indels 63; Gaps 10;

Qy 10 VLLLLQLLLAR-----SLDDGYKVEGKNAVLPSCYTLPTSGTLVPMCGKGF 58
Db 2 LLLLLPLMKRERAEAGTQSKLTMQSSVTVEGLCVHPCFSFSPSHGWTIPGVVAGY- 60
Qy 59 PWSOCTNELLRTDE-----RNVYQKSSRYQKGLDNKDVSLIKNVTLDHGT 108
Db 61 -WFR---EGANTDDAPVATNPPARAWEETRRDRFHLGLDHTKXCTLSIDARSDAGR 116
Qy 109 YCCCIQFPGI-MNDKLEKLDIKAKTTPAQTAGSDTASPRTLT----- 154
Db 117 YFFEMKSGIKMKNKHLRSLVNTALTPRPILPGLTESCPONTLCSSVWACEQGTTP 176
Qy 155 -----TERNGSETQTLVTLHNNGTKISTWDEKDSGETIRT--AIHIG 197
Db 177 MISWIGTSVSLDSDSTR--SSVLTILIPQPDHGTSLTC---QVTFPGASVTTKTYHLN 231
Qy 198 VGV-SAGITLALIG 211
Db 232 VSVPPQNLVMTVFG 246

```

```

RESULT 15
CD4_SAISC STANDARD; PRT; 457 AA.
ID_CD4_SAISC
AC Q29037;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Len-3).
GN CD4.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RA Tatsumi M., Hashimoto O.;
RU Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL
CC RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
CC -1- SUBUNIT: ASSOCIATES WITH P56-LCK (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

```

CC -----
DR EMBL; D86588; BAA131.1; -.
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:0005026; F: coreceptor activity; ISS.
DR GO; GO:0042289; F: MHC class II protein binding activity; ISS.
DR GO; GO:0006955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0045058; P: T-cell selection; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4_TCRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCRANTIGEN.
DR SMART; SM00406; IgV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 457 T-CELL SURFACE GLYCOPROTEIN CD4.
FT DOMAIN 26 395 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 396 417 POTENTIAL.
FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 126 202 IG-LIKE V-TYPE.
FT DOMAIN 203 316 IG-LIKE C2-TYPE 1.
FT DOMAIN 317 373 IG-LIKE C2-TYPE 2.
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 PALMITATE (BY SIMILARITY).
FT LIPID 421 421 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 457 AA; 50871 MM; 57EED6344005A015 CRC64;

```

```

Query Match 7.5%; Score 109.5; DB 1; Length 457;
Best Local Similarity 24.9%; Pred. No. 0.09;
Matches 46; Conservative 26; Mismatches 80; Indels 33; Gaps 8;

Qy 1 MFSGTLNCLVLLQLLLARSLDDGYKVEGKN---AYLPSCYTLPTSGTLVPMCGKGF 57
Db 1 MNGGIPRHLVLLQLLLPVNTHGKTVGKGEVVELPETSLKKN---VPRHWKTS- 56
Qy 58 CPWSQ-----CTNELRTDERNVYQKSSRYQKGLDNKDVSLIKNVTLDHGTCCR 112
Db 57 ---DQIKILGQNVFVRGSGSKLDRIDSK---KSSWDRGSFPLLDARIEDSETYICE 110
Qy 113 IQPGLMNDKLEKLDIKAKTTP-AQTAGSDT-----ASPRTLTERNGSETQ 163
Db 111 VE-----SKKEVELGVFGITANPDTHLLOGSLTTLLESPPGSSPSECTSPRGKRIR 164
Qy 164 TLVTL 168
Db 165 GRKTL 169

```

Search completed: November 22, 2003, 05:41:03  
 Job time : 16.5189 secs

THIS PAGE BLANK (USPTO)



QY 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 DB 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTWA 180  
 DB 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTWA 180  
 QY 181 DEIKSGEITRTAHIHIGVVSAGLTALIIIGVILIKWYCKKKKLSLITLANLPFG 240  
 DB 181 DEIKSGEITRTAHIHIGVVSAGLTALIIIGVILIKWYCKKKKLSLITLANLPFG 240  
 QY 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281  
 DB 241 LANAGAVRIRSEENITYTEENYEVENSNEYCYVNSQOPS 281

## RESULT 2

Q96K94 PRELIMINARY: PRT; 301 AA.  
 ID Q96K94  
 AC Q96K94;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein FLJ14428 (Putative kidney injury molecule-3).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabeura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;  
 RT "NEO human cDNA sequencing project";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hepatosoma;  
 RA Kuenen E.W., Ichimura T., Bonventre J.V.;  
 RT "A homolog to human kidney injury molecule-1 is expressed in hepatoma cells";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK027334; BAB55044.1; -;  
 DR EMBL; AY069944; BAB55401.1; -;  
 DR InterPro; IPR007110; IG-like.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 301 AA; 33351 MW; 4877B7D9EA0A4161 CRC64;

Query Match 60.1%; Score 883; DB 4; Length 301;  
 Best Local Similarity 63.7%; Pred. No. 1.1e-67;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFGSLTNCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCKGKGFPCW 60  
 DB 1 MESHLPFDVLLLLLLLRSSSEVEYRAEVGNAYLPCPTTPAAGNLVPCWKGKACPV 60  
 QY 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 DB 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTW 179  
 DB 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTW 179  
 QY 181 DEIKSGEITRTAHIHIGVVSAGLTALIIIGVILIKWYCKKKKLSLITLANLPFG 240  
 DB 181 DEIKSGEITRTAHIHIGVVSAGLTALIIIGVILIKWYCKKKKLSLITLANLPFG 240

DB 180 ANELRDSRLANDLRSGATIRIGIYGACGLALALIFGALIFKWSHSEKIQNLJL 239  
 QY 231 ITLANLPBGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 281  
 DB 240 ISLANLPBGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 291

## RESULT 3

Q8TD00 PRELIMINARY: PRT; 301 AA.  
 ID Q8TD00  
 AC Q8TD00;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus cellular receptor 2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Zhang W., Wan T., Li N., Cao X.;  
 RA "Novel human hepatitis A virus cellular receptor";  
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF251707; AAM19100.1; -;  
 DR Genew; HGNC:18437; HAVCR2.  
 DR InterPro; IPR003599; IG-  
 DR InterPro; IPR007110; IG-like.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Receptor.  
 SQ SEQUENCE 301 AA; 33363 MW; 7DAEA13F01D3587B CRC64;

Query Match 59.7%; Score 877; DB 4; Length 301;  
 Best Local Similarity 63.4%; Pred. No. 3.7e-67;  
 Matches 185; Conservative 26; Mismatches 69; Indels 12; Gaps 4;

QY 1 MFGSLTNCVLLLLLOLLARSLSDGYKVEGNAYLPCSYTLPTSGTLVPMCKGKGFPCW 60  
 DB 1 MESHLPFDVLLLLLLLRSSSEVEYRAEVGNAYLPCPTTPAAGNLVPCWKGKACPV 60  
 QY 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 DB 61 SGGTNEELRTDERNTYOKSSRYQLKGDINKGDSLLIKNTVLLDHDGTYCCRIQPFGLMN 120  
 QY 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTW 179  
 DB 121 DKLEIKLIDIKAKVTPAOTAHGDSSTTASPRITLTERNGSEOTVLTLLNNNGTKISTW 179  
 QY 180 ANELRDSRLANDLRSGATIRIGIYGACGLALALIFGALIFKWSHSEKIQNLJL 239  
 DB 180 ANELRDSRLANDLRSGATIRIGIYGACGLALALIFGALIFKWSHSEKIQNLJL 239  
 QY 231 ITLANLPBGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 281  
 DB 240 ISLANLPBGGLANAGAVRIRSEENITYTEENYEVENSNEYCYVNS-QOPS 291

## RESULT 4

Q8WM60 PRELIMINARY: PRT; 142 AA.  
 ID Q8WM60  
 AC Q8WM60;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Similar to hypothetical protein FLJ14428.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 SQ SEQUENCE 142 AA; 33363 MW; 7DAEA13F01D3587B CRC64;

RA Straubeberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC020843; AA020843.1; -.  
DR InterPro; IPR003599; IG\_-.  
DR InterPro; IPR007110; IG-like.  
DR SMART; SM00409; IG\_1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 142 AA; 16149 MW; E2C82D8B5A8C23E CRC64;

Query Match 29.3%; Score 430.5; DB 4; Length 142;  
Best Local Similarity 63.2%; Pred. No. 3e-29;  
Matches 86; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 1 MFSGLTNCVLLTLQLLARSLEDYKVEGKNAVLPSCSYTLPTSGLTVPMCKGKGFPCW 60  
DB 1 MFSGLTNCVLLTLQLLARSLEDYKVEGKNAVLPSCSYTLPTSGLTVPMCKGKGFPCW 60  
QY 61 SQTNELLRTERDENVYQKSSRYQKGLDKNGDVSLIKVTLDDHGTCCRIQPGIIMN 120  
DB 61 FEGCNVVLRTDERDVNVW-TSRVWLNGBFRKGDVSLTIENVTLTADSGIYCCRIQPGIIMN 119  
QY 121 DKLELKLIDIKAAVY 136  
DB 120 DEKNLKLVIKPGWT 135

## RESULT 5

OSVIM1 PRELIMINARY; PRT; 282 AA.  
ID OSVIM1  
AC OSVIM1;  
DT 01-MAR-2002 (TRENBLREL. 20, Created)  
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
DE TIM1.  
GN TIM1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DA/2; TISSUE=Spleen;  
RA McIntire J.J., Umetan S.E., Akbari O., Potter M., Barsh G.S.,  
RA Freeman G.J., Umetan D.T., Dekryff R.H.;  
RT "Apur. a major T cell regulatory locus that controls the development  
RT of airway hyperreactivity, cosegregates with variants in a novel gene  
RT family.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF399830; AA15775.1; -.  
DR MGD; MGI:2159680; Tmtd1.  
DR InterPro; IPR003599; IG\_-.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; IG\_1.  
DR SMART; SM00409; IG\_1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
SQ SEQUENCE 282 AA; 30968 MW; 7D30E0698F0AC5F CRC64;

Query Match 25.5%; Score 374; DB 11; Length 282;  
Best Local Similarity 34.9%; Pred. No. 5.3e-24;  
Matches 95; Conservative 42; Mismatches 105; Indels 30; Gaps 8;

QY 16 LLARSLEDYK---KVEGKNAVLPSCSYTLPTSGLTVPMCKGKGFPCW 72  
DB 12 LLARSLEDYKVEGKNAVLPSCSYTLPTSGLTVPMCKGKGFPCW 72  
QY 73 RNVYQKSSRYQKGLDKNGDVSLIKVTLDDHGTCCRIQPGIIMN 131  
DB 70 HRVYQKSSRYQKGLDKNGDVSLIKVTLDDHGTCCRIQPGIIMN 129  
QY 132 -----AAKVTAPQTAHGDSTASPT-----LTENSGSETQTLVTHANNNGTKI 176

DB 130 EIPTRPRRPTTRPATGRTTISTRTHVPTSTRVSTSPPTSTHTWTHKEDMNGVT 189  
QY 177 S---TWAD--EIDSGERT---AHIGVSGAGTLIALIGLILKWSCKKKLSL 228  
DB 190 SSGDTSMNTEALPPKPGQNPCKGFGVIGICIA-LILLVSTVAITRIYIMKRSASL 248  
QY 229 SLITLANLPPGLANAGAVRISEENIYITIE 260  
DB 249 SVAVFRSKIEALQNAVAVHRAEDNIYIED 280

## RESULT 6

OSV947 PRELIMINARY; PRT; 307 AA.  
ID OSV947  
AC OSV947;  
DT 01-JUN-1998 (TRENBLREL. 06, Created)  
DT 01-JUN-1998 (TRENBLREL. 06, Last sequence update)  
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
DE Kidney injury molecule-1 precursor (KIM-1).  
GN KIM-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;  
RX MEDLINE=98129827; PubMed=9461608;  
RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,  
RA Cate R.L., Sanicola M.;  
RT "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion  
RT molecule containing a novel immunoglobulin domain, is up-regulated in  
RT renal cells after injury";  
RL J. Biol. Chem. 273:4135-4142(1998).  
CC -1- FUNCTION: MAY ACT AS AN EPITHELIAL CELL ADHESION MOLECULE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN LIVER, SPLEEN AND  
CC NORMAL KIDNEY. LEVELS INCREASE IN THE POSTISCHEMIC KIDNEY WITH  
CC EXPRESSION FOUND IN REGENERATING PROXIMAL TUBULE EPITHELIAL CELLS.  
CC -1- INDUCTION: IN RENAL CELLS AFTER INJURY.  
CC -1- SIMILARITY: BELONGS TO THE IMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC ONE V-LIKE DOMAIN.  
DR EMBL; AF035963; AAC53546.1; -.  
DR InterPro; IPR003599; IG\_-.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; IG\_1.  
DR SMART; SM00409; IG\_1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
KW Transmembrane; Cell adhesion; Glycoprotein; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 307  
FT DOMAIN 22 235  
FT TRANSMEM 236 256  
FT DOMAIN 257 307  
FT DOMAIN 22 130  
FT DOMAIN 33 37  
FT DOMAIN 131 201  
FT DOMAIN 131 234  
FT DOMAIN 138 152  
FT REPEAT 138 142  
FT REPEAT 143 147  
FT REPEAT 148 152  
FT DISULFID 37 108  
FT CARBOHYD 206 206  
FT CARBOHYD 218 218  
SQ SEQUENCE 307 AA; 33963 MW; 736DIDD1F1549760 CRC64;

Query Match 24.9%; Score 365; DB 11; Length 307;  
Best Local Similarity 31.9%; Pred. No. 3.5e-23;  
Matches 94; Conservative 43; Mismatches 106; Indels 52; Gaps 9;

QY 14 LQLLARSLEDYKVE---EVGKNAVLPSCSYTLPTSGLTVPMCKGKGFPCW 70

[illegible]

RESULT 7			
ID	OSVIM2	PRELIMINARY;	PRT; 305 AA.
AC	OSVIM2		
DT	01-MAR-2002 (TREMBlrel. 20,	Created)	
DT	01-MAR-2002 (TREMBlrel. 20,	Last sequence update)	
DT	01-MAR-2003 (TREMBlrel. 23,	Last annotation update)	
DE	TIM1.		
GN	TIMD1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/c; TISSUE=Spleen;		
RA	McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,		
RA	Frieman G.J., Umetsu D.T., Dextruyf R.H.;		
RT	"Tapr, a major T cell regulatory locus that controls the development		
RT	of airway hyperreactivity, cosegregates with variants in a novel gene		
RT	family.";		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL, AF398929; AA3574.1; -.		
DR	MED; MGI:2155680; Timd1.		
DR	InterPro; IPR003599; Ig.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	Pfam; PF00047; Ig; 1.		
DR	SMART; SM00409; Ig; 1.		
DR	PROSITE; PS50835; Ig_LIKE; 1.		
QO	SEQUENCE 305 AA; 33391 MW;		8F4EA38627FE85FB CRC64;

Query Match	24.8%	Score 364.5;	DB 11;	Length 305;
Best Local Similarity	31.2%	Pred. No. 3.8e-23;		
Matches	92;	Conservative.	39;	Mismatches 111;
			Indels	53;
			Gaps	7;

Qy	16	LLIARSLLEGY---	KVEVEKNALPESYTLPTSGTLVPMCMGKGFCSWSCOTHELARTDE	72
Db	12	ILILBGTUDSYVEKGVASHPTLPCTYS--	TYRGLTTTCMGKGCQPSACQNTLLMTNG	69
Qy	73	RNVYQKSSRYOLKGDINKGVSLLIKNVTLDDHGYTCRIQFPGILNNDKLELKDIK-	1311	
Db	70	HRVTQKSSRYMLKGHISEGDVSLTIENSVESSGGLCYCCREVEIPGFWMFNQKVTFSLQVKP	1299	
Qy	132	-----AAKVPACQAHGDSFTAS-----	-PPTLTERNGSTQTLVYLLHANNNGTKI--	1766
Db	130	EIPTRPPTTRPTTRPATGRPTTISRSRTHVTPFSINVESTPPTSTHTWTWTHKEPTEFCP	1899	
Qy	177	-----STWDEIKDGEETIRT-----	AHIGVGSAGLT	2050
Db	190	HETTAEVNGIPBSHTPDMNGVLTYSGGDTMSNHTHEALPPGKRPQKNPTTKGFVGLGICIAA-LL	2488	
Qy	206	LALIIGVLLIKWYSCCKKKLSSILSLITLNLPPQGLANNAGVIRISEENIYTTIEE	260	

Db 249 LLLVSTVAITRYILMKRKSASLSVAFRVSKIEALQNAAVVHSHRAEDNIYIVED 303

ID	PRELIMINARY;	PRT;	359 AA.
043656			

DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus cellular receptor 1.  
 DE HRAVCR-1.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

R#	SEQUENCE FROM N.A.
RC	TISSUE=Liver;
RX	MEDLINE=98325180; PubMed=9658108;
RA	Fingelstrock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;
RT	"The human homolog of HAVc-1 codes for a hepatitis A virus cellular receptor.";
RT	J. Virol. 72:6621-6628(1998).
RL	EMBL, AF043724, AAC39862.1; -
DR	InterPro: IPR003599; IG_
DR	InterPro: IPR007110; IG-like.
DR	InterPro: IPR003006; IG_MHC.
DR	Pfam: PF00047; IG_1.
DR	SMART, SM00409; IG_1.
DR	PROSITE, PS00835; IG_LIKE; 1.
DR	PROSITE, PS00290; IG_MHC; 1.
KM	Receptor.
SQ	SEQUENCE 359 AA, 38704 MW, C207FECS62DC6A2CA CRG4;

Query Match	22.44;	Score 328.5;	DB 4;	Length 359;
-------------	--------	--------------	-------	-------------

Matches 102; Conservative 49; Mismatches 93; Indels 123; Gaps 11;

```

QY      10 VLLILQLLARSLEDEGKLV--EVGNNAVLPSPYLPTSGTLVPMCMWGKGFCSPISSQCTNEL 67
        ||| ||| : ||| ||| : ||| : ||| :
DB      6 VIISLIHLADSVAGSVKVGEGAPSVTLPCHY----SGAVTSMCMWRGSCSLFTCQNGI 61

```

```

68 LRTERNVTVYQKSSRYQLKGDILNKGSVSLIKNVTLDHGTGCRIGPGLMNDKLELR 12
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 VMTGTGHTVYRKDFRYKLGLDLSRRDVSLTIENTAVSDSGVYCCRVHGWFNDMKITVS 12

```

```

QY      128 LDI+AAKVT-----13
      |:| |
DB      122 LELI+PKVTTTPIVTTVPVTVTRKSTTVPTTTVPPTTVPTTMSIPTTTTVPTMTVST 183

```

```

QY      137 -----PAQTAGDSTTAS-----PRTL----- 15
          |      :  || |
Db      182 TTSTVPTTSTPTTTSVPVTTTSTFVPRPMLRQNHHPVATSPSSPQBAETHPTTLOGAI 24

```

```

Qy 154 -----TTERNGSEQTLYTLHNNNGTKISWADKLKDSGEIIRT--AIHIGVG 199
      || : | : : || || : : || : : || : : ||
Db 242 RREPSSPLSYTTIDGNDYITESSDGLMNNQQLF-----LEHSLLTANTTKCIYAGVC 290

```

Qy 200 VSAGITLALIGVLLIKWYCKKKKLSLTLANIPGGGLANAGAVRIRSEENITYTE 25  
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 297 ISV-LVLALLLGVIIAKKYFF-KKEVQQLS-VSFSSIQIKALONAVEKEVQAEDNITY-IE 35

Qy	260	ENVYEVE	266
		:: :	
Db	353	NSLYATD	359

RESULT 9	
Q8VBW0	
ID Q8VBW0	PRELIMINARY;
AC Q8VBW0;	PRT; 305 AA

RESULT 9	
Q8VBW0	
ID Q8VBW0	PRELIMINARY;
AC Q8VBW0;	PRT; 305 AA.



DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE TIM2.  
 GN TIM2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 NC NCI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2, and BALB/c; TISSUE=Spleen;  
 RA McInliffe J.J., Umetsu S.E., Akbari O., Potter M., Barch G.S.,  
 RA Freeman G.J., Umetsu D.T., Dekruyff R.H.;  
 RT "Tadp, a major T cell regulatory locus that controls the development  
 RT of airway hyperreactivity, cosegregates with variants in a novel gene  
 RT family."  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF39828; AAL35773.1; -  
 DR EMBL; AF39828; AAL35772.1; -  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR SMART; PS50835; IG\_LIKE; 1.  
 SQ SEQUENCE 305 AA; 33519 MW; 68E69370FE23C9D0 CRC64;

Query Match 22.2%; Score 326; DB 11; Length 305;  
 Best Local Similarity 29.3%; Pred. No. 7.7e-20;  
 Matches 87; Conservative 49; Mismatches 117; Indels 44; Gaps 8;

QY 14 LQLLARLEEDGKYK--VGKNAVLPCTYLTPTSGTLVPMCKGFCFMSQCTNELTLD 71  
 DB 11 LILLPGAVESHATAVQGLAGHPVTLPCITYSTHGLG-IVPMCKGEGECHSYCIRSLMTWN 69  
 QY 72 ERNVTYQKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIQFPGLANM--DKKLELKD 129  
 DB 70 GYVTHQNSRYOLKGNISSEGNVSLTIENTVGDGPRCCVETIGAHFPMYLEVPE 129  
 QY 130 IK-----AAKVTPAQTAHGDSTASPTLTERRNGS----- 161  
 DB 130 ISTSPPTPTATGPRPTTISTRTSTHVTSTRTSTSTSPATETERYKPEATTFYPDQTAA 189  
 QY 162 -TQTLVTL---HNNNGTKISTWADI-----KSGETIRRAIHGVGSAGLTALIG 211  
 DB 190 VTEITLPSIPADWHNTVTSDDPDNDTEVIPPQKQKNLNGFYGISIAA-LILMLLS 248  
 QY 212 VLILKMWCKKKKSLSLITLANLPQGLANAGAVIRSENIYTIENYEVENS 268  
 DB 249 TMVITRYVMKRSLSLFAFPISKIGASPKKVERTRCEDOVIIIDTTPPEES 305

## RESULT 10

Q96D42 PRELIMINARY; PRT; 364 AA.

AC Q96D42.  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NC NCI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strauberg R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013325; AAH13325.1; -  
 DR Genew; HGNC:17866; HAVCR1.

DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81P1 CRC64;

Query Match 22.2%; Score 326; DB 4; Length 364;  
 Best Local Similarity 27.4%; Pred. No. 9.8e-20;  
 Matches 102; Conservative 49; Mismatches 93; Indels 128; Gaps 11;

QY 10 VLLILQLLARLEEDGKYK--EVGKNAVLPCTYLTPTSGTLVPMCKGFCFMSQCTNEL 67  
 DB 6 VILSLIHLADSVAGSVKVGAGPSVTLPCY---SGAVTSMCKNGSCSLFTCCNGI 61  
 QY 68 LRTDERNVTYQKSSRYOLKGDINKGVSLLIKNVTLDHGTCCRIQFPGLANDKKLEL 127  
 DB 62 VMTNGTHTVTRKDRKYLGLSLSRDVSLTIENTAVDSGVCCRVHERGMFNDMKITVS 121  
 QY 128 LDIKAAYT----- 136  
 DB 122 LEIVPKVTTPIVTVPTVTRSTVPTTTPVMTTVPTTTPPTTMSIPTTVLVT 181  
 QY 137 -----PAQTAHGDSTAS-----PRT 152  
 DB 182 MTSVSTTSVPTTTSIPTTSPVTTVSTFVPMPLPRONHEPVATSPSSQOPAEHTPTT 241  
 QY 153 L-----TTERNGSETQTLVTLHNNNGTKISTWADIKSGETIRT--AI 194  
 DB 242 LQGAIRREPTSPXYSTTDNDVTSSDDLANNNOTQLF-----LEHSLTLANTTGI 296  
 QY 195 HIGVGSAGLTALIIIGVLIKMWCKKKKSLSLITLANLPQGLANAGAVIRSEEN 254  
 DB 297 YAGCISY-LVLLALLGVIIAKKYF--KKEVQQLS-VSFSLOIKALONAVEKEVQADN 353  
 QY 255 IYTIENYEVYE 266  
 DB 354 IY-IENSLVATD 364

## RESULT 11

Q8R183 PRELIMINARY; PRT; 305 AA.

AC Q8R183.  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical 33.5 kDa protein.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strauberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025096; AAH25096.1; -  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 305 AA; 33504 MW; 410SD75D99F541AB CRC64;

Query Match 21.7%; Score 318; DB 11; Length 305;  
 Best Local Similarity 28.3%; Pred. No. 3.8e-19;  
 Matches 84; Conservative 48; Mismatches 121; Indels 44; Gaps 7;



FT SIGNAL 1 17 POTENTIAL.  
 SQ SEQUENCE 451 AA; 48774 MM; 5D395C5455AA4332 CRC64;  
 Query Match 18.2%; Score 267.5; DB 6; Length 451;  
 Best Local Similarity 21.7%; Pred. No. 1.4e-14;  
 Matches 97; Conservative 51; Mismatches 96; Indels 203; Gaps 11;

QY 7 LNCVLLQLLLARSLDEGYVE--VGKNAVLPSCSYTLPTSGTLVPMCMGKGFPMWSQCT 64  
 DB 3 LQVVLISLIHLADSVADVNDVAGISITLPCRY----NGAITSNMGHNGTCSVFSCP 58  
 QY 65 NELIRTERNVYQKSSRYQLKGLNKGDVSLIKNVTLDHGYCCRIQPPGLNDKCL 124  
 DB 59 DGIWVTNGTHTVYKRETRYKLLGNLSRBDVSLTIANAVSDSGIYCCRVKSGWFNDMKI 118  
 QY 125 ELKIDIKAKAYT----- 136  
 DB 119 TISLKGPPRVTPIVRTVSTTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTL 178  
 QY 137 ----- 136  
 DB 179 LPTTTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTLP 238  
 QY 137 ----- 136  
 DB 239 TTTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPT 298  
 QY 137 -----PAQT-----AHGDTASP-RITITERNGSETOVLVTHNNGTKIST 178  
 DB 299 PVATSPSPPOAETHPTVLGATRTQPTSSPLSYSTDSDIVESSDGLMNNQTOISP 358  
 QY 179 MADEIKSGETRT--AIHIGVSGAGTLALIGVLLIKWYSCCKKKLSLITLNL 236  
 DB 359 -----EHSPOKVNTEGIVAGVCISV-LVLAIVGVIKAYFF-KKEIQOLS-VSFSNH 410  
 QY 237 PPGIANGAVRIRSENIYTIENNV 263  
 DB 411 QFKTLQNAVKEVHADNIV-1ENNLY 436

RESULT 15  
 ID 046597 PRELIMINARY; PRT; 474 AA.  
 AC 046597;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus cellular receptor 1 long form (Hepatitis A virus  
 cellular receptor 1 short form).  
 GN HAVCR-1.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RA Feigelestock D., Thompson P., Mattoo P., Kaplan G.G.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF043447; AAC39772.1; -;  
 DR EMBL; AF043446; AAC39771.1; -;  
 DR InterPro; IPR003599; IG-  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 474 AA; 51088 MM; 472DF987CA220524 CRC64;

QY 7 LNCVLLQLLLARSLDEGYVE--VGKNAVLPSCSYTLPTSGTLVPMCMGKGFPMWSQCT 64  
 DB 8 LQVVLISLIHLADSVADVNDVAGISITLPCRY----NGAITSNMGHNGTCSVFSCP 63  
 QY 65 NELIRTERNVYQKSSRYQLKGLNKGDVSLIKNVTLDHGYCCRIQPPGLNDKCL 124  
 DB 64 DGIWVTNGTHTVYKRETRYKLLGNLSRBDVSLTIANAVSDSGIYCCRVKSGWFNDMKI 123  
 QY 125 ELKIDIKAKAY-TPAQTAHGDTASPRTLITERNGSETOVLVTL 168  
 DB 124 TISLKGPPRVTPIVRTVSTTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTL 168

Search completed: November 22, 2003, 05:43:21  
 Job time : 67.5584 secs

Query Match 18.1%; Score 265.5; DB 6; Length 474;  
 Best Local Similarity 38.2%; Pred. No. 2.2e-14;  
 Matches 63; Conservative 24; Mismatches 71; Indels 7; Gaps 3;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 01:46:35 ; Search time 58.9588 Seconds  
(without alignments)  
810.341 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587  
Sequence: 1 MFSLPDPVCVLLLLLLTR.....CVSSRQPSQPLGCRFAMP 301

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	301	17	AAW01049
2	1587	100.0	301	21	AAV97058
3	1587	100.0	301	22	AAW50223
4	1587	100.0	301	22	AAU14409
5	1587	100.0	301	22	AAV93838
6	1587	100.0	301	22	AAV81518
7	1587	100.0	301	22	AAV51104
8	1587	100.0	301	22	AAV59169
9	1587	100.0	301	23	ABG32775

10	1587	100.0	301	24	ABP70445	Amino acid sequenc
11	1587	100.0	301	24	ABG73718	Human TH1-associat
12	1581	99.6	301	24	ABP70444	Amino acid sequenc
13	1580	99.6	301	22	AAU14173	Human novel protei
14	1580	99.6	301	23	ABV93036	Human polypeptide
15	1449	91.3	401	24	ABJ37916	NOVX protein sequ
16	940	59.2	185	20	AAV02369	Polypeptide identi
17	885	55.8	281	24	ABP70435	Amino acid sequenc
18	883	55.6	281	17	AAW01047	Product of 200 gen
19	883	55.6	281	21	AAV97056	Murine T helper ce
20	883	55.6	281	22	AAV50225	Murine 200 gene pr
21	883	55.6	281	22	AAV81511	Murine TH1 specif
22	883	55.6	281	22	AAV50696	Murine 200 gene pr
23	883	55.6	281	23	AAV59162	Murine 200 gene pr
24	883	55.6	281	23	ABG32767	Murine TH1 specif
25	883	55.6	281	24	ABP70434	Amino acid sequenc
26	883	55.6	281	24	ABG73716	Murine TH1-associ
27	718	45.2	142	21	AAV19730	Human SECX Clone 4
28	672	42.3	125	20	AAV12263	Human 5' EST seque
29	336	21.2	307	19	AAW38334	Rat kidney injury
30	333	21.0	282	19	AAV70431	Amino acid sequenc
31	323.5	20.4	305	24	ABP70430	Amino acid sequenc
32	322	20.3	359	24	ABR48174	Human bladder canc
33	322	20.3	359	24	ABP70438	Amino acid sequenc
34	322	20.3	359	24	ABP70439	Amino acid sequenc
35	322	20.3	359	24	ABP70441	Amino acid sequenc
36	322	20.3	359	24	AAV06680	KIM-1 related prot
37	319.5	20.1	364	24	ABP70442	Amino acid sequenc
38	319	20.1	364	24	ABP70440	Amino acid sequenc
39	317.5	20.0	364	24	ABP70443	Amino acid sequenc
40	273.5	17.2	305	24	ABP70432	Amino acid sequenc
41	273.5	17.2	305	24	ABP70433	Amino acid sequenc
42	267	16.8	334	19	AAV38336	Human kidney injur
43	267	16.8	334	19	AAV26679	KIM-1 related prot
44	256	16.1	451	17	AAV28803	Hepatitis A virus
45	250.5	15.8	345	24	ABP70436	Amino acid sequenc

#### ALIGNMENTS

RESULT 1	AAW01049	standard; Protein: 301 AA.
ID	AAW01049	
AC	AAW01049;	
XX	29-DEC-1996	(first entry)
DT		
XX	Product of 200 gene differentially expressed in T helper cells.	
XX		
KW	T helper cell; TH cell; T-cell; T-lymphocyte; 200 gene;	
KW	differential expression; immune disorder; multiple sclerosis;	
KW	asthma; lepromatous leprosy; diagnosis; therapy; receptor.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..20
FT		/label= sig_peptide
FT	Domain	21..200
FT		/label= Extracellular_domain
FT	Domain	201..224
FT		/label= Transmembrane_domain
FT	Domain	225..301
FT		/label= Cytoplasmic_domain
XX		
XX	W09627603-A1.	
XX		
XX	12-SEP-1996.	
XX		
XX	01-MAR-1996;	96WO-US02798.
XX		

PR	07-JUN-1995;	95US-0487748.
PR	03-MAR-1995;	95US-0398633.
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Levinson DA;	
XX		
DR	WPI; 1996-433404/43.	
XX	N-PSDB; AAT38267.	
PT	Genes and their products differentially expressed in T helper cells	
PT	- useful in diagnosis and treatment of immune disorders, e.g.	
PT	multiple sclerosis, asthma, lepromatous leprosy, etc.	
XX		
PS	Claim 10; Fig 24; 218pp; English.	
CC	The product (AAW01049) of novel human gene 200 (AAT38267) is a novel	
CC	cell surface receptor of the Ig superfamily class. Gene 200	
CC	expression is many-fold higher in TH1 than in TH2 subpopulations.	
CC	Modulation of the 200 gene product may ameliorate a range of	
CC	T-cell-related disorders. Soluble gene 200 products (e.g. fusions	
CC	to immunoglobulins) can be produced that increase the blood	
CC	half-life of the product. Transgenic animals expressing the 200	
CC	gene product are useful models of TH cell subpopulation-related	
CC	disorders. The murine homologue (AAW01047) of the gene 200 product	
CC	has also been identified.	
XX		
SQ	Sequence 301 AA;	
Query Match	100.0%; Score 1587; DB 17; Length 301;	
Best Local Similarity	100.0%; Pred. No. 8.8e-144;	
Matches 301; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MESHLPPOCVLLILLLLLTTRSESEYRABVGONAYLPCTYTPAARGNLVPCWKGACPV 60	
DB	1 MESHLPFCVLLILLLLLTTRSESEYRABVGONAYLPCTYTPAARGNLVPCWKGACPV 60	
QY	61 FECGNVAVRTERDVNYWTSRYWLNQDFPRKGVSLTIENVTLADSGICRIQIPGINND 120	
DB	61 FECGNVAVRTERDVNYWTSRYWLNQDFPRKGVSLTIENVTLADSGICRIQIPGINND 120	
QY	121 EKENLKIVYKPAKVPAPTLQDFPAAPPRMLTTRGHGPAETQYLSLPDINTQISTLA 180	
DB	121 EKENLKIVYKPAKVPAPTLQDFPAAPPRMLTTRGHGPAETQYLSLPDINTQISTLA 180	
QY	181 NELRDSRLANDRDSGATIRIGYAGICAGALALIFGALIFQWYSHSKKIONLSLI 240	
DB	181 NELRDSRLANDRDSGATIRIGYAGICAGALALIFGALIFQWYSHSKKIONLSLI 240	
QY	241 SLANLPPSGLANAANVAGIRSEENITYIENNVYEEPEPNYYCYVSSROQSPGLGCRPAM 300	
DB	241 SLANLPPSGLANAANVAGIRSEENITYIENNVYEEPEPNYYCYVSSROQSPGLGCRPAM 300	
QY	301 P 301	
DB	301 P 301	
RESULT 2		
AAV97058		
ID	AAV97058 standard; Protein; 301 AA.	
XX		
AC	AAV97058;	
XX		
DT	31-OCT-2000 (first entry)	
XX		
DE	Human T helper cell differentially expressed gene 200 product.	
XX		
KW	T helper cell; differential expression; 200 gene; immunomodulator;	
KW	anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;	
KW	thymometric; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;	
XX	protozoacide; lymphocyte; modulator; gene therapy.	

OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Domain
FT	Domain
FT	Domain
FT	Domain
FT	Domain
FT	Domain
PN	US6084083-A.
XX	
PD	04-JUL-2000.
XX	
XX	
PF	28-MAR-1997;
XX	
PR	01-MAR-1996;
PR	03-MAR-1995;
PR	07-JUN-1995;
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Levinson DA;
XX	
DR	WPI; 2000-464385/40.
XX	
XX	N-P9DB; AHA51904; AHA51905.
PT	New isolated human 200 gene products or polypeptides, useful for treating and diagnosing immune disorders, especially T helper lymphocyte-related disorders
PT	
PS	Claim 1; Fig 24A-D; 107bp; English.
XX	
CC	Genes which are differentially expressed within and among T helper (TH) cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations, can be used diagnostically or as targets for therapeutic intervention.
CC	The polypeptides are useful for treating and diagnosing of immune disorders, especially T lymphocyte-related disorders. These disorders include chronic inflammatory diseases and disorders (e.g. Crohn's disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or Grave's disease), or atopic conditions (e.g. asthma and allergy, including allergic rhinitis or food allergies). Also included are certain pathogen susceptibilities (e.g. leishmaniasis), and viral (e.g. HIV) or bacterial (e.g. tuberculosis or leptomatous leprosy) infections.
CC	
XX	
SQ	Sequence 301 AA;
	Query Match 100.0%; Score 1587; DB 21; Length 301;
	Beet Local Similarity 100.0%; Pred. No. 8.8e-144; Indels 0; Gaps 0;
	Matches 301; Conservative 0; Mismatches 0;
OY	1 MFSLPEPCVALLLLLTSSSEVEYAEGVQNALPCEFYTPAAPGMLVPCKMGKACPV 60
Db	1 MFSLPEPCVALLLLLTSSSEVEYAEGVQNALPCEFYTPAAPGMLVPCKMGKACPV 60
OY	61 FEGGNVLRTDERDVNTWSRYWLNGDFRKDVSLLTENVTTLADSGIYCRRIQIPGIIND 120
Db	61 FEGGNVLRTDERDVNTWSRYWLNGDFRKDVSLLTENVTTLADSGIYCRRIQIPGIIND 120
OY	121 EKFNKLIVTKAKTTPAPTLORDSTAAFPRLTLTRHGCPATOTLGSIPDINTLOISTLA 180
Db	121 EKFNKLIVTKAKTTPAPTLORDSTAAFPRLTLTRHGCPATOTLGSIPDINTLOISTLA 180
OY	181 NELRDSRLANDLRDSGATIRIGIYIGAGICGLALAIIFGLAIFPKWYSKEKIQNLSLI 240
Db	181 NELRDSRLANDLRDSGATIRIGIYIGAGICGLALAIIFGLAIFPKWYSKEKIQNLSLI 240
OY	241 SLANLPESGLANAAYAEGRSEENYYTTEENVYEVEEPNEYYCYVSRQOPSOPLGCRFAM 300
XX	

Db 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSRQPSQPLGCRPFAM 300  
 QY 301 P 301  
 Db 301 P 301

RESULT 3  
 ID AAM50223 standard; Protein; 301 AA.  
 AC AAM50223;  
 DT 07-JAN-2002 (first entry)

DE Human 200 gene product, preferentially expressed in TH1 cells.  
 XX  
 KW Human; 200 gene; T helper; T lymphocyte; T cell; TH1;  
 KW receptor; differential expression; immune disorder; psoriasis;  
 KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
 KW antiporiatic; diagnosis; therapy.

OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Signal\_peptide  
 FT Protein 21..301  
 FT /label= Mature\_protein  
 FT Domain 21..200  
 FT /label= Extracellular\_domain  
 FT Domain 201..224  
 FT /label= Transmembrane\_domain  
 FT Domain 225..301  
 FT /label= Cytoplasmic\_domain

XX US6288218-B1.  
 XX  
 PN 11-SEP-2001.  
 PD 25-SEP-1997; 97US-0937399.  
 PF 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX (LEVI/) LEVINSON D A.  
 PA Levinson DA;  
 PI WPI; 2001-647189/74.  
 DR N-PSDB; AAI70254; AAI70281.  
 DR  
 XX  
 PT Detecting 200 gene expression in a sample, useful for treating and  
 PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
 PT comprising detecting the presence of a 200 gene product or an RNA  
 PT encoding the 200 gene product -  
 XX  
 PS Claim 1(a); Fig 24A-B; 108pp; English.

XX The present sequence is that of the protein product of the human  
 CC 200 gene (see AAI70254). It is a receptor of the Ig superfamily  
 CC class. The 200 gene is preferentially expressed in mature, fully  
 CC differentiated T helper subpopulation TH1 cells relative to  
 CC subpopulation TH2 cells. The gene can be used diagnostically or as  
 CC a target for therapeutic intervention for the treatment of immune  
 CC disorders. A claimed method for diagnosing a TH cell subpopulation  
 CC related immune disorder involves detecting the level of a 200 gene  
 CC product, or an RNA encoding it, so that if the level differs from  
 CC that in a control sample, the disorder is diagnosed. The gene  
 CC product detected may be the present amino acid sequence, or  
 CC amino acids 1-20, 1-200, 1-224, 30-128, 21-200, 21-301,  
 CC 201-224, 201-301 or 224-301 of it. The immune disorder is

CC especially a TH1 cell subpopulation-related immune disorder, such  
 CC as multiple sclerosis, psoriasis or insulin-dependent diabetes  
 CC (claimed). In addition to the 200 gene, the invention provides  
 CC other genes that are differentially expressed within and among  
 CC TH cells and TH cell subpopulations and which can be used in  
 CC methods for the diagnosis, prognosis, evaluation and treatment of  
 CC TH cell subpopulation-related disorders, for the identification of  
 CC subjects exhibiting a predisposition to such conditions, for  
 CC monitoring patients undergoing clinical evaluation for the  
 CC treatment of such disorders, and for monitoring the efficacy of  
 CC compounds used in clinical trials. Other immune disorders that can  
 CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
 CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,  
 CC Grave's disease, contact dermatitis, graft rejection, graft versus  
 CC host disease, sarcoidosis, atopic conditions, asthma, allergy,  
 CC allergic rhinitis, food allergy, eosinophilia, conjunctivitis,  
 CC glomerular nephritis, helminthic infection (e.g. leishmaniasis),  
 CC viral infection (e.g. HIV), and bacterial infection (e.g.  
 CC tuberculosis and lepromatous leprosy).

XX Sequence 301 AA;  
 SQ

Query Match 100.0%; Score 1587; DB 22; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPPDCVLLILLTLTSSSEVEYRAVGQNNYLPCFTTPAPGNLVPVCKGKACP 60  
 Db 1 MFSLPPDCVLLILLTLTSSSEVEYRAVGQNNYLPCFTTPAPGNLVPVCKGKACP 60

QY 61 FECSNVVLRDERDVNWTGRWYMLNGDFRKGVSLTEENTVTLADSGIYCCRIQPGIMND 120  
 Db 61 FECSNVVLRDERDVNWTGRWYMLNGDFRKGVSLTEENTVTLADSGIYCCRIQPGIMND 120

QY 121 EKENLKVIRKAVYTPAFTLQDFTPAAPPMLTTRGHGPAETQGLSPDINTQISTLA 180  
 Db 121 EKENLKVIRKAVYTPAFTLQDFTPAAPPMLTTRGHGPAETQGLSPDINTQISTLA 180

QY 181 NEFDSRLANDLRSGATIRIGIYGICGALALIPGLIRKWSHSEKIQNSLI 240  
 Db 181 NEFDSRLANDLRSGATIRIGIYGICGALALIPGLIRKWSHSEKIQNSLI 240

QY 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSRQPSQPLGCRPFAM 300  
 Db 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSRQPSQPLGCRPFAM 300

QY 301 P 301  
 Db 301 P 301

RESULT 4  
 ID AAU14409 standard; Protein; 301 AA.  
 AC AAU14409;  
 DT 24-OCT-2001 (first entry)

XX Human novel protein #280.  
 DE  
 XX  
 KW Human; novel protein; antianemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;  
 KW anticonvulsant; antiarrhythmic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.

OS Homo sapiens.  
 XX  
 PN WO200155437-A2.

```

PD 02-AUG-2001.
XX
XX 25-JAN-2001; 2001MO-US02623.
XX
XX 25-JAN-2000; 2000US-0491404.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dymnac RT;
XX
XX WPI; 2001-451939/48.
XX
XX N-PSDB; AAS22714.
XX
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
XX nervous system disorders, and for regenerating bone and cartilage.
XX
XX Example 4; Page 804-805; 894pp; English.
XX
XX The invention relates to polynucleotides encoding novel human
XX proteins or their active domains. The polypeptides, polynucleotides and
XX antibodies raised against the polypeptides are used in a method of
XX treatment of a mammal and prevention of disorders caused by the aberrant
XX protein expression or activity. The polypeptides can be used as
XX molecular weight markers, food supplements, and in antibody production.
XX The polypeptides are used to identify compounds which bind to the
XX polypeptides. Polynucleotides of the invention are used as probes and
XX primers, for sequencing, for chromosome or gene mapping, in the
XX production of recombinant proteins, and in generating anti-sense DNA or
XX RNA and in gene therapy. Polypeptides of the invention can be used to
XX target drugs to a tumour, in assays to determine biological activity, to
XX raise antibodies/ elicit an immune response, to determine quantitative
XX protein levels, as tissue markers, and to isolate receptors or ligands.
XX Polypeptides of the invention may also be useful in treating platelet
XX disorders, stem cell disorders, wound healing, treating burns, promoting
XX ligament and/or nerve tissue, regenerating bone, cartilage, tendon,
XX the proliferation, differentiation and survival of stem cells, as a
XX contraceptive, treating osteoporosis and osteoarthritis, anaemia,
XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX fungal infection or from autoimmunity, cancer, allergy, asthma,
XX graft-versus-host disease, eczema, haemophilia, thrombosis,
XX anti-inflammatory diseases, nervous system disorders, and infection.
XX The present sequence represents a protein of the invention.
XX
XX Sequence 301 AA;
SQ
Query Match 100.0%; Score 1587; DB 22; Length 301;
Best Local Similarity 100.0%; Pred. No. 8.8e-144;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFSHLPDCVLLLLLLLRSSSEVERAEGVGNAYLPCFTTPAAGNLVPCMGKACPV 60
DB 1 MFSHLPDCVLLLLLLLRSSSEVERAEGVGNAYLPCFTTPAAGNLVPCMGKACPV 60
QY 61 FEGGNVVLRTDESDVNYMTSRVYLANDPRKGDVSLTIENTVADSGITCORTIPIGIMND 120
DB 61 FEGGNVVLRTDESDVNYMTSRVYLANDPRKGDVSLTIENTVADSGITCORTIPIGIMND 120
QY 121 EKFNKLVIKPAKVTPAPTLQDRFTAAPRMLTTRHGAPETQTLGSLPDINTQISTLA 180
DB 121 EKFNKLVIKPAKVTPAPTLQDRFTAAPRMLTTRHGAPETQTLGSLPDINTQISTLA 180
QY 181 NEHRDSRLANDLDSGNTIRIGIYAGACAGLALLIFGALIFKYSHSKEXIQWLSLI 240
DB 181 NEHRDSRLANDLDSGNTIRIGIYAGACAGLALLIFGALIFKYSHSKEXIQWLSLI 240
QY 241 SLANLPSPGLANVAAGIRSEENIYITEENVYEEVEENENYCVVSRROOSPGLGRFPM 300
DB 241 SLANLPSPGLANVAAGIRSEENIYITEENVYEEVEENENYCVVSRROOSPGLGRFPM 300
QY 301 P 301
DB 301 P 301

```

```

RESULT 5
AAB93838
ID AAB93838 standard; Protein; 301 AA.
XX
XX AAB93838;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human protein sequence SEQ ID NO:13669.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
OS
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Mashikawa T, Hayashi K, Saito K, Yamamoto J;
XX Iehli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 13669; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 301 AA;
SQ
Query Match 100.0%; Score 1587; DB 22; Length 301;
Best Local Similarity 100.0%; Pred. No. 8.8e-144;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFSHLPDCVLLLLLLLRSSSEVERAEGVGNAYLPCFTTPAAGNLVPCMGKACPV 60

```



Db 1 MESHLPDCCVLLLLLLLITRSSEVEYRAEYVGNAYLPCEYTPAARGNLVPVCMGKACP 60  
QY 61 FECCNVVLRTERDVNVTWTSRYWLNQDFRKGDVSLTIENVTADSGIYCCRIQIPGIMND 120  
Db 61 FECCNVVLRTERDVNVTWTSRYWLNQDFRKGDVSLTIENVTADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIRPAKVTPTPTQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
Db 121 EKFNKLVIRPAKVTPTPTQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYAGICAGLALALIFGALIFKWTSHSXEKIQNLSLI 240  
Db 181 NELRDSRLANDLRDSGATIRIGIYAGICAGLALALIFGALIFKWTSHSXEKIQNLSLI 240  
QY 241 SLANLPPSGLANNAVAEGIRSEENIYTIENVEVEEENVEYCYVSSROQPSQPLGCRPFAM 300  
Db 241 SLANLPPSGLANNAVAEGIRSEENIYTIENVEVEEENVEYCYVSSROQPSQPLGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 6  
AAB81518  
ID AAB81518 standard; Protein; 301 AA.  
XX AC AAB81518;  
XX 18-JUN-2001 (first entry)  
XX DE Human TH1 specific 200 gene product.  
XX KM Human: T helper cell; TH cell; TH1; TH2; immunomodulator;  
KW anti-inflammatory; antiallergic; dermatological; antiviral;  
KW antibacterial; T helper lymphocyte modulator; gene therapy;  
KW TH specific gene; 200 gene; immune disorder; inflammation;  
KW infection.  
XX OS Homo sapiens.  
XX PN US6204371-B1.  
XX PD 20-MAR-2001.  
XX PF 01-MAR-1996; 96US-0609583.  
XX PR 03-MAR-1995; 95US-0398633.  
XX PR 07-JUN-1995; 95US-0487748.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Levinson DA;  
XX DR WPI; 2001-272703/28.  
XX DR N-PSDB; AAF82616.  
XX PT New murine or human 200 genes and their corresponding polypeptides,  
PT useful for treating or diagnosing immune disorders, especially T helper  
PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies  
XX PS Claim 5; Fig 24; 109pp; English.  
XX CC The present sequence is encoded by the human 200 gene, which is  
CC expressed at higher levels in T helper (TH)1 cells than in TH2 cells.  
CC The invention relates to an isolated nucleic acid molecule, which  
CC comprises the full length murine 200 gene or full length human 200  
CC gene nucleotide sequence. The nucleic acids are useful for treating  
CC or diagnosing immune disorders, especially T helper  
CC lymphocyte-related disorders, e.g. inflammatory diseases (e.g.  
CC Crohn's disease), multiple sclerosis, Grave's disease, contact  
CC dermatitis, psoriasis, asthma and allergies, or certain viral

CC (e.g. HIV) or bacterial (e.g. tuberculosis) infections.  
XX Sequence 301 AA;  
Query Match 100.0%; Score 1587; DB 22; Length 301;  
Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESHLPDCCVLLLLLLLITRSSEVEYRAEYVGNAYLPCEYTPAARGNLVPVCMGKACP 60  
Db 1 MESHLPDCCVLLLLLLLITRSSEVEYRAEYVGNAYLPCEYTPAARGNLVPVCMGKACP 60  
QY 61 FECCNVVLRTERDVNVTWTSRYWLNQDFRKGDVSLTIENVTADSGIYCCRIQIPGIMND 120  
Db 61 FECCNVVLRTERDVNVTWTSRYWLNQDFRKGDVSLTIENVTADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIRPAKVTPTPTQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
Db 121 EKFNKLVIRPAKVTPTPTQORDFTAAFPRLMTTRGHGPAETQTLGSLPDINTLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATIRIGIYAGICAGLALALIFGALIFKWTSHSXEKIQNLSLI 240  
Db 181 NELRDSRLANDLRDSGATIRIGIYAGICAGLALALIFGALIFKWTSHSXEKIQNLSLI 240  
QY 241 SLANLPPSGLANNAVAEGIRSEENIYTIENVEVEEENVEYCYVSSROQPSQPLGCRPFAM 300  
Db 241 SLANLPPSGLANNAVAEGIRSEENIYTIENVEVEEENVEYCYVSSROQPSQPLGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 7  
AAB51104  
ID AAB51104 standard; Protein; 301 AA.  
XX AC AAB51104;  
XX DT 20-MAR-2001 (first entry)  
XX DE Human 200 gene product amino acid sequence SEQ ID NO:24.  
XX KM Treatment; diagnosis; immune disorder; mast cell related disorder;  
KW T-helper lymphocyte-related disorder; ischemic disorder;  
KW identification; vasodilator; cardiac; antianginal; angina pectoris;  
KW ischaemic renal disease; myocardial ischaemia; myocardial infarction;  
KW cortical infarction; ischaemic injury; kidney transplant.  
XX OS Homo sapiens.  
XX PN WO200073498-A1.  
XX PD 07-DEC-2000.  
XX PF 31-MAY-2000; 2000WO-US14986.  
XX PR 02-JUN-1999; 99US-0324986.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Levinson DA, Lloyd CM, McCarthy SA;  
XX DR WPI; 2001-016510/02.  
XX DR N-PSDB; AAC90993.  
XX PT Ameliorating a symptom of an ischemic disorder or injury in a mammal  
PT e.g. ischemic renal disease or myocardial ischemia, by administering a  
PT 200 gene product (S1), a nucleic acid encoding (S1) or an antibody  
PT directed against (S1) -  
XX PS Claim 10; Fig 24; 309pp; English.

CC The present invention describes a method for ameliorating a symptom of  
CC an ischemic disorder or injury in a mammal. The method comprises  
CC administering a 200 gene product, a nucleic acid encoding (SI) or an  
CC antibody directed against (SI). The method is useful for treating a  
CC symptom of an ischemic disorder such as ischemic renal disease or  
CC myocardial infarction (such as angina pectoris), myocardial or cortical  
CC infarction. The method is also useful for treating a symptom of an  
CC ischemic injury occurring due to transplantation of a kidney. The  
CC present sequence is used in the exemplification of the present  
CC invention.

XX Sequence 301 AA;

XX Query Match 100.0%; Score 1587; DB 22; Length 301;

XX Best Local Similarity 100.0%; Pred. No. 8.8e-144;

XX Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLILLILLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCMGKACPV 60

DB 1 MESHLPFDCVLLILLILLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCMGKACPV 60

QY 61 FECGNVLTDERDVNYWTSRYWLNQDFRKGVSLTIENVTLADSGIYCCRIQIPGIMND 120

DB 61 FECGNVLTDERDVNYWTSRYWLNQDFRKGVSLTIENVTLADSGIYCCRIQIPGIMND 120

QY 121 EKFNKLKLVTKPAKVPAPTLQDPFTAAFPRLTTRHGPAEQTGSLPDINLTQISTLA 180

DB 121 EKFNKLKLVTKPAKVPAPTLQDPFTAAFPRLTTRHGPAEQTGSLPDINLTQISTLA 180

QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKKIQNLSLI 240

DB 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKKIQNLSLI 240

QY 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEBPENYCYVSSROQPSQPLGCRFAM 300

DB 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEBPENYCYVSSROQPSQPLGCRFAM 300

QY 301 P 301

DB 301 P 301

RESULT 8  
AAB59169 standard; protein; 301 AA.

XX ID AAB59169;

XX AC AAB59169;

XX DT 21-MAR-2001 (first entry)

XX DE Human 200 gene protein.

XX KW Cysteine protease; immune disorder; T lymphocyte; Crohn's;

XX KW arthritis; diabetes; multiple sclerosis; viral infection; bacterial;

XX KW HIV.

XX OS Homo sapiens.

XX PN US6156887-A.

XX PD 05-DEC-2000.

XX PF 03-OCT-1997; 97US-0939729.

XX PR 01-MAR-1996; 96US-0609583.

XX PR 03-MAR-1995; 95US-0398633.

XX PR 07-JUN-1995; 95US-0487748.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Levinson DA;

XX DR WPI; 2001-101473/11.

XX Novel polypeptide exhibiting cysteine protease activity, useful for  
PT treating and diagnosing immune disorders, especially T  
PT lymphocyte-related disorders, e.g. Crohn's disease, multiple sclerosis,  
PT graft versus host disease or allergies -  
XX Examples; Fig 24; 107pp; English.

CC The present invention relates to a novel cysteine protease. The  
CC protein of the invention is useful for treating and diagnosing immune  
CC disorders, especially T lymphocyte-related disorders. In particular,  
CC the polypeptide is useful for treating or diagnosing T helper (TH) cell  
CC or TH cell subpopulation-related disorders. These disorders include  
CC Crohn's disease, reactive arthritis, Lyme disease, insulin-dependent  
CC diabetes, organ-specific autoimmunity, multiple sclerosis, Hashimoto's  
CC thyroiditis, Grave's disease, contact dermatitis, psoriasis, graft  
CC rejection, graft versus host disease, sarcoidosis, atopic (e.g. asthma  
CC or allergy), eosinophilia, conjunctivitis, glomerular nephritis, or  
CC helminthic (e.g. leishmaniasis), viral (e.g. HIV (human  
CC immunodeficiency virus)) or bacterial (e.g. tuberculosis or lepromatous  
CC leprosy) infections.

XX Sequence 301 AA;

XX Query Match 100.0%; Score 1587; DB 22; Length 301;

XX Best Local Similarity 100.0%; Pred. No. 8.8e-144;

XX Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLILLILLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCMGKACPV 60

DB 1 MESHLPFDCVLLILLILLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCMGKACPV 60

QY 61 FECGNVLTDERDVNYWTSRYWLNQDFRKGVSLTIENVTLADSGIYCCRIQIPGIMND 120

DB 61 FECGNVLTDERDVNYWTSRYWLNQDFRKGVSLTIENVTLADSGIYCCRIQIPGIMND 120

QY 121 EKFNKLKLVTKPAKVPAPTLQDPFTAAFPRLTTRHGPAEQTGSLPDINLTQISTLA 180

DB 121 EKFNKLKLVTKPAKVPAPTLQDPFTAAFPRLTTRHGPAEQTGSLPDINLTQISTLA 180

QY 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKKIQNLSLI 240

DB 181 NELRDSRLANDLRDSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKKIQNLSLI 240

QY 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEBPENYCYVSSROQPSQPLGCRFAM 300

DB 241 SLANLPSPGLANAVAGIRSEENIYTIENNVYVEBPENYCYVSSROQPSQPLGCRFAM 300

QY 301 P 301

DB 301 P 301

RESULT 9  
ABG32775 standard; protein; 301 AA.

XX ID ABG32775;

XX AC ABG32775;

XX DT 15-NOV-2002 (first entry)

XX DE Human TH specific 200 gene protein.

XX KW TH1; TH2; ss; antiinflammatory; antidiabetic; antithyroid; atopy;

XX KW antiallergic; antiallergic; virucide; antiarthritic; dermatological;

XX KW antipsoriatic; nephrotoxic; immunosuppressive; immune response;

XX KW asthma; allergy; allergic rhinitis; viral infection; thyroiditis;

XX KW inflammatory disease; Crohn's disease; arthritis; diabetes; human;

XX KW dermatitis; psoriasis; glomerular nephritis; autoimmunity; gene;

XX KW graft rejection.

XX OS Homo sapiens.

EN US6414117-B1.  
 XX 02-JUL-2002.  
 PD 12-MAY-1999; 99US-0310367.  
 XX 01-MAR-1996; 96US-0609583.  
 PR 28-MAR-1997; 97US-0829525.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX  
 PI Levinson DA;  
 XX  
 DR MPI; 2002-641576/69.  
 DR N-PSDB; ABS53322.  
 XX  
 PT Novel polypeptides which are differentially expressed within and among  
 PT T helper cells and cell populations, useful for treating immune  
 PT disorders, especially T helper cell subpopulation-related disorders -  
 XX  
 PS Example; Fig 24; 109pp; English.  
 XX  
 CC This invention relates to a novel isolated polypeptide which is a T  
 CC helper (Th1) cell subpopulation specific gene product. The invention  
 CC also discloses other Th1 or Th2 specific gene products. The proteins of  
 CC the invention may have antiinflammatory, antidiabetic, antithyroid,  
 CC antiasthmatic, antiallergic, virocidic, antiautistic, dermatological,  
 CC antiproliferative, nephrotropic and immunosuppressive activities and may be  
 CC used as a regulator of the immune response. The proteins of the  
 CC invention are also useful to reduce the level of Th2 cell activity for  
 CC treating Th1 cell subpopulation-related disorders including atopic  
 CC conditions such as asthma and allergy including allergic rhinitis, the  
 CC effects of pathogen, including viral infection, chronic inflammatory  
 CC diseases such as Crohn's disease, arthritis, diabetes, thyroiditis,  
 CC dermatitis, psoriasis, glomerular nephritis, organ-specific  
 CC autoimmunity, graft rejection and graft versus host disease. The  
 CC present sequence represents the human 200 gene protein of the invention.  
 CC  
 XX  
 SO Sequence 301 AA;  
 Query Match 100.0%; Score 1587; DB 23; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLLTRSEVEYRAVGNAYLPCTYPAPAGNLPVPCWKGACPV 60  
 DB 1 MFSLPFCVLLLLLLTRSEVEYRAVGNAYLPCTYPAPAGNLPVPCWKGACPV 60  
 QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 DB 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 181 NELRDSRLANDIRDSGATIRIGIYAGICAGLALALIFGALIFKWSHSEKIONLSLI 240  
 DB 181 NELRDSRLANDIRDSGATIRIGIYAGICAGLALALIFGALIFKWSHSEKIONLSLI 240  
 QY 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEPEPEYCYVSSROOPSGPCRRAM 300  
 DB 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEPEPEYCYVSSROOPSGPCRRAM 300  
 QY 301 P 301  
 DB 301 P 301  
 RESULT 10  
 ABP70445

ID ABP70445 standard; Protein; 301 AA.  
 XX  
 AC ABP70445;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of human TIM-3 allele 2.  
 XX  
 DE T cell immunoglobulin domain; mucin domain; TIM-1; TIM-2; TIM-3;  
 XX TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 XX myelodysplastic syndrome; airway hyperactivity; cancer; asthma;  
 XX allergic T cell response; autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 140  
 FT /note="encoded by CCG"  
 XX  
 PD WO200302722-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 PF 01-JUL-2002; 2002WO-US20890.  
 XX  
 PR 29-JUN-2001; 2001US-302344P.  
 XX  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
 DR MPI; 2003-210268/20.  
 DR N-PSDB; ABZ68339.  
 XX  
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 PS Claim 10; Page 90; 94pp; English.  
 XX  
 CC The present sequence is a human T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved IGV and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperactivity and allergic T cell responses, and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
 CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
 CC  
 XX  
 SO Sequence 301 AA;  
 Query Match 100.0%; Score 1587; DB 24; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSLPFCVLLLLLLTRSEVEYRAVGNAYLPCTYPAPAGNLPVPCWKGACPV 60  
 DB 1 MFSLPFCVLLLLLLTRSEVEYRAVGNAYLPCTYPAPAGNLPVPCWKGACPV 60  
 QY 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 DB 61 FECGNVLRTERDVNWTSRWYLNQDPRKGVSLTIENVTLADSGIYCCRIQIPGINND 120  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 QY 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180  
 DB 121 EKFNKLVIRKPAKVTAPTLQDFTAAFPRLTTRGHPATQTLGSLPDINLTQISTLA 180

QY 181 NELRDSRLANDLRDGSATIRIGIYAGICAGLALALIFGALIFKMYSHSKKXIONLSLI 240  
 DB 181 NELRDSRLANDLRDGSATIRIGIYAGICAGLALALIFGALIFKMYSHSKKXIONLSLI 240  
 QY 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRPFAM 300  
 DB 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRPFAM 300  
 QY 301 P 301  
 DB 301 P 301

RESULT 11  
 ABG73718  
 ID ABG73718 standard; Protein; 301 AA.  
 AC ABG73718;  
 XX  
 DT 09-APR-2003 (first entry)  
 XX  
 DE Human TH1-associated 200 protein SEQ ID 24.  
 KW T-cell receptor; TH; T helper cell; 103 gene; TH2 cell; human;  
 KW TH2 cell marker; TH1 cell; protozoacide; antibacterial; virucide;  
 KW immunosuppressive; antiinflammatory; antirheumatic; antidiabetic;  
 KW neuroprotective; dermatological; antichyroid; antipruritic; helminthic;  
 KW nephrotropic; antiallergic; CD8 agonist; CD4 agonist;  
 KW interleukin agonist; bacterial; viral infection; immune disorder;  
 KW Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;  
 KW Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;  
 KW graft rejection; graft versus host disease; asthma; glomerulonephritis;  
 KW allergy; gene therapy; TH cell subpopulation.  
 XX  
 XX Homo sapiens.  
 OS  
 PN US6455685-B1.  
 PD 24-SEP-2002.  
 XX  
 XX 27-FEB-1998; 98US-0032337.  
 PF  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 PR 01-MAR-1996; 96US-0609583.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 PI Levinson DA;  
 XX  
 DR WPI, 2003-066247/06.  
 DR N-PSDB; ABQ77040, ABQ77041.  
 XX  
 PT Identifying a test compound that binds to a 103 gene product or is a  
 PT TH2 cell marker, useful for the identification and therapeutic use of  
 PT compounds as treatments of helminthic, bacterial and viral infections,  
 PT and immune disorders -  
 XX  
 XX Example 9; Figure 24A-C; 131p; English.  
 PS  
 CC This invention describes a novel method for identifying a test compound  
 CC that binds to a 103 gene product or is a TH2 cell marker. The method  
 CC comprises contacting a test compound with an immobilised 103 gene  
 CC product, removing unbound test compound or separating the complex from  
 CC the reaction mixture, and detecting the complex. Identifying a test  
 CC compound that binds to a 103 gene product alternatively comprises: (a)  
 CC contacting a test compound with a cell engineered to express a 103 gene  
 CC product or co-expressing a 103 gene product and a test compound in a  
 CC cell; and (b) removing unbound 103 gene product. Identifying a test  
 CC compound that is a TH2 cell marker further comprises detecting a complex,  
 CC where contacting the test compound with a TH2 and TH1 cell for the test  
 CC compound to bind either cell and detecting binding where it indicates the  
 CC test compound as a TH2 cell marker. The products described in the

CC invention have protozoacide, antibacterial, virucide, immunosuppressive,  
 CC antiinflammatory, antirheumatic, antidiabetic, neuroprotective,  
 CC dermatological, antichyroid, antipruritic, nephrotropic, antiallergic;  
 CC and antiallergic activity and can act as CD8, CD4 and interleukin  
 CC agonists. The methods and compositions of the present invention are  
 CC useful for the identification and therapeutic use of compounds as  
 CC treatments of helminthic, bacterial and viral infections and immune  
 CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
 CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
 CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
 CC asthma, allergy and glomerulonephritis. They can also be used for gene  
 CC therapy, for the diagnostic evaluation and prognosis of TH cell  
 CC subpopulation-related disorders, identification of subjects exhibiting a  
 CC predisposition to such conditions, monitoring undergoing clinical  
 CC evaluation and efficacy for the treatment of the disorders. This sequence  
 CC represents a T-Helper cell associated polypeptide described in the  
 CC disclosure of the invention.  
 XX  
 SQ Sequence 301 AA;  
 XX  
 Query Match 100.0%; Score 1587; DB 24; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-144;  
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MESHPDPCVLLLLLLLRSEVEYRAVGONAVLPCYTPPAAGNLVPCMGKACPV 60  
 DB 1 MESHLPDPCVLLLLLLLRSEVEYRAVGONAVLPCYTPPAAGNLVPCMGKACPV 60  
 QY 61 FEGCNVVAJRTDENVNWTSRWMLNGDFPKGDSLTIEENVTLADSGIYCCR1QIPQIMND 120  
 DB 61 FEGCNVVAJRTDENVNWTSRWMLNGDFPKGDSLTIEENVTLADSGIYCCR1QIPQIMND 120  
 QY 121 EKFNKLVIYKPAKVPAPLQDFTAAPPRMLTTGHPAEQTQGSJPDINLTQISTLA 180  
 DB 121 EKFNKLVIYKPAKVPAPLQDFTAAPPRMLTTGHPAEQTQGSJPDINLTQISTLA 180  
 QY 181 NELRDSRLANDLRDGSATIRIGIYAGICAGLALALIFGALIFKMYSHSKKXIONLSLI 240  
 DB 181 NELRDSRLANDLRDGSATIRIGIYAGICAGLALALIFGALIFKMYSHSKKXIONLSLI 240  
 QY 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRPFAM 300  
 DB 241 SLANLPPSGLANAVAGIRSEENIYTIENVEVEEPNEYYCVSSROQPSOPLGCRPFAM 300  
 QY 301 P 301  
 DB 301 P 301

RESULT 12  
 ABP70444  
 ID ABP70444 standard; Protein; 301 AA.  
 AC ABP70444;  
 XX  
 DT 22-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of human TIM-3 allele 1.  
 DE  
 XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
 KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
 KW allergic T cell response; autoimmune disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 140  
 FT /note="encoded by CDS"  
 PN W02003002722-A2.  
 XX  
 XX 09-JAN-2003.

```
XX 01-JUN-2002; 2002WO-US20890.
PF
XX
XX 29-JUN-2001; 2001US-302344P.
PR
XX
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
PA
XX
XX McIntire JJ, Dekruff RH, Umetsu DT, Freeman GJ, Kuchroo V;
PI
XX MPI: 2003-210268/20.
DR
XX N-PSDB; ABZ68338.
DR
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain
PT and mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease -
XX
XX Claim 10; Page 89; 94pp; English.
XX
XX The present sequence is a human T cell immunoglobulin domain and
CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,
CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The
CC TIM gene family is located within a region of human chromosome 5 that
CC is commonly deleted in malignancies and myelodysplastic syndrome.
CC Variants of TIM-1 and TIM-3 are associated with susceptibility to
CC atway hyperreactivity and allergic T cell responses and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1
CC cells preferentially express TIM-3, while Th2 cells preferentially
CC express TIM-1. TIM polypeptides and polynucleotides are useful for
CC treating cancer, asthma, allergies, eczema or autoimmune diseases.
XX
XX Sequence 301 AA:
SQ
Query Match 99.6%; Score 1581; DB 24; Length 301;
Best Local Similarity 99.7%; Pred. No. 3.3e-143;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MFSHLPDCCVLLLLLLLITRSEVYRAEYVQGNALPCFYTPAAPGNLVPVCGKACPV 60
DB 1 MFSHLPDCCVLLLLLLLITRSEVYRAEYVQGNALPCFYTPAAPGNLVPVCGKACPV 60
OY 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTADSGIYCCRIQIPGINND 120
DB 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTADSGIYCCRIQIPGINND 120
OY 121 EKFNKLVIKPAKVTAPFTLQRFDAFPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180
DB 121 EKFNKLVIKPAKVTAPFTLQRFDAFPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180
OY 181 NELNDSRLANDLRDSGATIRIGIYAGICAGLALALFGALIFWYSHSEKIONLSLI 240
DB 181 NELNDSRLANDLRDSGATIRIGIYAGICAGLALALFGALIFWYSHSEKIONLSLI 240
OY 241 SLANLPESGLANAAVEGIRSEENITTIENNYEVEPEPEYCYVSSROOPSOLGCRPAM 300
DB 241 SLANLPESGLANAAVEGIRSEENITTIENNYEVEPEPEYCYVSSROOPSOLGCRPAM 300
OY 301 P 301
DB 301 P 301
RESULT 13
AAU14173
ID AAU14173 standard; Protein; 301 AA.
XX
XX AAU14173;
AC
XX 24-OCT-2001 (first entry)
DT
XX Human novel protein #44.
DE
```

```
XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerrary; noctropic;
KW anticonvulsant; antichratic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
XX Homo sapiens.
OS
XX WO200155437-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 25-JAN-2001; 2001WO-US02623.
PF
XX 25-JAN-2000; 2000US-0491404.
PR
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dzmanac RT;
PI
XX MPI: 2001-451939/48.
DR
XX N-PSDB; AAS22478.
DR
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
XX
XX Example 4; Page 550; 894pp; English.
XX
XX The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/elicitor an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence represents a protein of the invention.
XX
XX Sequence 301 AA:
SQ
Query Match 99.6%; Score 1580; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 4.1e-143;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MFSHLPDCCVLLLLLLLITRSEVYRAEYVQGNALPCFYTPAAPGNLVPVCGKACPV 60
DB 1 MFSHLPDCCVLLLLLLLITRSEVYRAEYVQGNALPCFYTPAAPGNLVPVCGKACPV 60
OY 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTADSGIYCCRIQIPGINND 120
DB 61 FECGNVLTARDERDVNYWTSRYWLNQDPRKGDVSLITENVTADSGIYCCRIQIPGINND 120
OY 121 EKFNKLVIKPAKVTAPFTLQRFDAFPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180
DB 121 EKFNKLVIKPAKVTAPFTLQRFDAFPRMLTTRGHGPAETQTLGSLPDINLTQISTLA 180
```

Db 121 EKENLKLVIKPAKVTAPPTLQDRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NEIRDRLANDLRDSCATTIRIGIYGAGTCAGLALALIFGALLFKWYSHSKKXIONLSLI 240  
Db 181 NEIRDRLANDLRDSCATTIRIGIYGAGTCAGLALALIFGALLFKWYSHSKKXIONLSLI 240  
QY 241 SLANLPPSGLANAVAEGRSEENITYTEENVEVEEENVEYCYVSSROOPSQPLGCRFAM 300  
Db 241 SLANLPPSGLANAVAEGRSEENITYTEENVEVEEENVEYCYVSSROOPSQPLGCRFAM 300  
QY 301 P 301  
Db 301 P 301  
RESULT 14  
ID ABB90396 standard; Protein; 301 AA.  
XX ABB90396;  
AC ABB90396;  
DT 24-MAY-2002 (first entry)  
XX Human polypeptide SEQ ID NO 2772.  
DE Human polypeptide SEQ ID NO 2772.  
XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;  
KW antidiabetic; hepatotropic; antidiabetic; antinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX Homo sapiens.  
OS Homo sapiens.  
PN WO200190304-A2.  
XX 29-NOV-2001.  
PD 29-NOV-2001.  
XX 18-MAY-2001; 2001MO-US16450.  
PF 18-MAY-2001; 2001MO-US16450.  
XX 19-MAY-2000; 2000US-205515P.  
PR 19-MAY-2000; 2000US-205515P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX Birse CE, Rosen CA;  
PI Birse CE, Rosen CA;  
XX WPI, 2002-122018/16.  
DR N-PSDB; ABL90805.  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive, and  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX Claim 11; SEQ ID NO 2772; 2081pp + Sequence Listing; English.  
XX The invention relates to novel genes (ABL9449-ABL90853) and proteins  
CC (ABB90404-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases such as cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WFO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 301 AA:  
Query Match 99.6%; Score 1580; DB 23; Length 301;  
Best Local Similarity 99.7%; Pred. No. 4, 1e-143;  
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSHLPEFDCVLLILLILLTRSSSEVEYRAVGONALPCPYTAAAGNLVPCWKGCACV 60  
Db 1 MFSHLPEFDCVLLILLILLTRSSSEVEYRAVGONALPCPYTAAAGNLVPCWKGCACV 60  
QY 61 FEGCNVLTERTDERDNYWTSRYWLNQDFPKGDSVLTIEVTLADSGIYCCRIQIPEINMD 120  
Db 61 FEGCNVLTERTDERDNYWTSRYWLNQDFPKGDSVLTIEVTLADSGIYCCRIQIPEINMD 120  
QY 121 EKENLKLVIKPAKVTAPPTLQDRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
Db 121 EKENLKLVIKPAKVTAPPTLQDRDFTAAFPRLMTTRGHGPAETQTLGSLPDINLTQISTLA 180  
QY 181 NEIRDRLANDLRDSCATTIRIGIYGAGTCAGLALALIFGALLFKWYSHSKKXIONLSLI 240  
Db 181 NEIRDRLANDLRDSCATTIRIGIYGAGTCAGLALALIFGALLFKWYSHSKKXIONLSLI 240  
QY 241 SLANLPPSGLANAVAEGRSEENITYTEENVEVEEENVEYCYVSSROOPSQPLGCRFAM 300  
Db 241 SLANLPPSGLANAVAEGRSEENITYTEENVEVEEENVEYCYVSSROOPSQPLGCRFAM 300  
QY 301 P 301  
Db 301 P 301  
RESULT 15  
ID ABB90396 standard; Protein; 401 AA.  
XX ABB90396;  
AC ABB90396;  
DT 22-MAY-2003 (first entry)  
XX 22-MAY-2003 (first entry)  
XX NOVA protein sequence SEQ ID NO 77.  
XX Hepatotropic; immunosuppressive; cardiant; hypertensive; tranquilizer;  
KW vulnerary; virucide; antibacterial; prozoocide; fungicide; nocotropic;  
KW antiparasitic; neuroprotective; cerebroprotective; antiparkinsonian;  
KW anticonvulsant; antidiabetic; analgesic; dermatological; keratolytic;  
KW antiseborrheic; antihemetic; anticholelithic; antinflammatory; anti-HIV;  
KW cytostatic; antiaesthetic; antipsoriatic; hypotensive; osteopathic;  
KW antitumor; anorectic; antidiabetic; antiallergic; haemostatic;  
KW neuroleptic; antidepressant; antinfertility; NOVA; human disease;  
KW NOVA-associated disorder; trauma; viral; bacterial; fungal; protozoal;  
KW parasitic infection; Alzheimer's disease; stroke; forensic biology;  
KW immunogen; non-human transgenic animal; gene therapy.  
XX Unidentified.  
OS Unidentified.  
PN WO200281517-A2.  
XX 17-OCT-2002.  
PD 17-OCT-2002.  
XX 22-JAN-2002; 2002MO-US02064.  
PF 22-JAN-2002; 2002MO-US02064.  
XX 19-JAN-2001; 2001US-262892P.  
PR 19-JAN-2001; 2001US-262892P.  
XX 23-JAN-2001; 2001US-263598P.  
PR 23-JAN-2001; 2001US-263598P.  
XX 25-JAN-2001; 2001US-264179P.  
PR 25-JAN-2001; 2001US-264179P.  
XX 26-JAN-2001; 2001US-264439P.  
PR 26-JAN-2001; 2001US-264439P.  
XX 30-JAN-2001; 2001US-263515P.  
PR 30-JAN-2001; 2001US-263515P.  
XX 02-MAR-2001; 2001US-272870P.  
PR 02-MAR-2001; 2001US-272870P.  
XX 14-MAR-2001; 2001US-275927P.  
PR 14-MAR-2001; 2001US-275927P.  
XX 15-MAR-2001; 2001US-276449P.  
PR 15-MAR-2001; 2001US-276449P.

20-MAR-2001; 2001US-277358P.  
 23-MAR-2001; 2001US-278151P.  
 29-MAR-2001; 2001US-279857P.  
 20-APR-2001; 2001US-285140P.  
 20-APR-2001; 2001US-285141P.  
 30-APR-2001; 2001US-287484P.  
 17-MAY-2001; 2001US-291701P.  
 08-JUN-2001; 2001US-296960P.  
 10-JUL-2001; 2001US-304353P.  
 10-JUL-2001; 2001US-304355P.  
 12-JUL-2001; 2001US-304886P.  
 09-AUG-2001; 2001US-311289P.  
 13-AUG-2001; 2001US-311975P.  
 16-AUG-2001; 2001US-312937P.  
 18-OCT-2001; 2001US-330227P.  
 29-NOV-2001; 2001US-334198P.  
 (CURA-) CURAGEN CORP.  
 Decristofaro MF, Padigaru M, Miller C, Tchernev V, Zhong H;  
 Zhong M, Anderson D, Ballinger R, Gerlach V, Spytek KA;  
 Rastelli L, Kehuda R, Guo X, Zernhusen B, Andrew D, Mezes P;  
 Patcuajan M, Burgess CE, Eissen A, Wolenc A, Baumgartner U;  
 Shumke RA, Gusev V, Vernet CM, Taupier RJ, Pena C, Shenoy S;  
 Li L, Casman S, Bollog F, Fernandes E, Smithson G, Malyankar U;  
 Tallon B, Liu X;  
 MPI: 2003-058504/05.  
 N-ESDB; AET33381.  
 New polypeptides, designated as NOVX, useful for diagnosing and  
 treating infections, neurological diseases, cancer, allergy, and bone,  
 immunological, skin, renal, brain, muscle and autoimmune disorders -  
 Claim 1; Page 209; 672pp; English.  
 The invention relates to a novel isolated polypeptide, designated NOVX  
 (NOV1 - 33), consisting of a mature form of one of 61 sequences, given  
 in the specification, or its variant, where amino acid residue(s) in the  
 variant differ from the mature form, provided that the variant differs  
 in not more than 15 % of the amino acids from the sequence of the mature  
 form. The NOVX polypeptides, nucleic acids encoding the polypeptides, and  
 an antibody to the polypeptides, are useful for treating or preventing a  
 NOVX-associated disorder in humans and for treating a syndrome associated  
 with a human disease (NOVX-associated disorder). NOVX polypeptides and  
 the encoding nucleic acids, are useful for determining the presence of or  
 predisposition to a disease associated with altered levels of NOVX  
 polypeptide and polynucleotide, by measuring the level of polypeptide  
 expression or the amount of nucleic acid from a mammal and comparing it  
 with another mammal not having or not predisposed to the disease. NOVX  
 polypeptide is also useful for identifying an agent that binds to NOVX  
 and a cell expressing NOVX is useful for identifying an agent that  
 modulates the expression or activity of NOVX. The antibodies and a  
 polypeptide having 95 % sequence identity to NOVX polypeptide are useful  
 for treating a pathological state in a mammal. The antibodies are also  
 useful for determining the presence or amount of NOVX in a sample. NOVX  
 polypeptides, polynucleotides and antibodies specific for the  
 polypeptides are useful for treating or preventing disorders or syndromes  
 including trauma, viral, bacterial, fungal, protozoal, and parasitic  
 infections. They can also treat disorders such as e.g., Alzheimer's  
 disease or a stroke. The NOVX encoding nucleic acids are useful for  
 expressing the NOVX proteins, to detect NOVX mRNA, or a genetic lesion in  
 a NOVX gene and to modulate NOVX activity. NOVX sequences are also useful  
 for identifying a cell or tissue type in a biological sample, to amplify  
 DNA sequences from very small biological samples such as tissues e.g.  
 hair or skin or body fluids in forensic biology and as primers and probes  
 for use in identifying and/or cloning NOVX homologues in other cell  
 types. The NOVX proteins are useful as an immunogen to generate  
 antibodies which are useful for diagnostically monitoring protein levels  
 and modulating NOVX activity. Cells comprising NOVX nucleic acids are  
 useful for producing non-human transgenic animals which are useful for  
 studying the function and/or activity of NOVX protein and for identifying  
 and/or evaluating modulators of NOVX protein activity. The NOVX nucleic

CC acids can be used in gene therapy. This sequence represents a NOVX  
 CC protein of the invention.  
 XX  
 SO Sequence 401 AA;  
 Query Match 91.3%; Score 1449; DB 24; Length 401;  
 Best Local Similarity 72.3%; Pred. No. 2,4e-130;  
 Matches 290; Conservative 3; Mismatches 8; Indels 100; Gaps 2;  
 QY 1 MFSLPDCVLLLLLT-----17  
 DB 1 MFSLPDCVLLLLLTTLFSRVHRIQPPSPRAVFLKVRONGVKQSDNAFVSGGD 60  
 QY 18 -----LTR-----20  
 DB 61 RTGKDELSTICLAQTRTFPLVKGKMGSSFLPLIDMACSPFEMELPHSPLBSWLSFLYAG 120  
 QY 21 SSEVEYRAEVQNNAYLPCFYTPAAPGNLVPCWKGACPVPECGNVVLRTERDVTWTS 80  
 DB 121 SSEVEYRAEVQNNAYLPCFYTPAAPGNLVPCWKGACPVPECGNVVLRTERDVTWTS 180  
 QY 81 RYWLNGDFRRGDSLTLENTVTLADSGTYCCRIQIPIGIMDEKFKLVIPAKVTPA 140  
 DB 181 RYWLNGDFRRGDSLTLENTVTLADSGTYCCRIQIPIGIMDEKFKLVIPAKVTPA 240  
 QY 141 QRDFTAFPRMLTTRGHGPAETQTLGSLPDINLTQISTLANELRDSRLANDRDSGATIR 200  
 DB 241 QRDFTAFPRMLTTRGHGPHDGSCLSPDVRLTQISTLANELRDSRLANDRDSGATIR 300  
 QY 201 IGIYAGICAGLALIFGALIFKWSHSEKTIQNLISLANLPSSGLANAAVEGIRS 260  
 DB 301 IGIYAGICAGLALIFGALIFKWSHSEKTIQNLISLANLPSSGLANAAVEGIRS 360  
 QY 261 EENITTEENYEEVEENYCYVSSROPSQPCPEFAMP 301  
 DB 361 EENITTEENYEEVEENYCYVSSROPSQPCPEFAMP 401  
 Search completed: November 22, 2003, 05:40:21  
 Job time : 59.9588 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:38:42 ; Search time 19.6529 Seconds  
(without alignments)  
648.024 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MFSHLPDVCVLLLLLLTR.....CYVSSROOPSQPLCRFPAMP 301

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	301	US-08-829-525-24	Sequence 24, Appli
2	1587	100.0	301	US-08-609-583A-24	Sequence 24, Appli
3	1587	100.0	301	US-08-937-399-24	Sequence 24, Appli
4	1587	100.0	301	US-09-310-367-24	Sequence 24, Appli
5	1587	100.0	301	US-09-033-337-24	Sequence 24, Appli
6	1587	100.0	301	US-09-464-231-24	Sequence 24, Appli
7	883	55.6	281	US-08-487-748A-9	Sequence 9, Appli
8	883	55.6	281	US-08-487-748A-10	Sequence 10, Appli
9	883	55.6	281	US-08-480-070C-10	Sequence 10, Appli
10	883	55.6	281	US-08-829-525-10	Sequence 10, Appli
11	883	55.6	281	US-08-937-399-10	Sequence 10, Appli
12	883	55.6	281	US-09-310-367-10	Sequence 10, Appli
13	883	55.6	281	US-09-464-231-10	Sequence 10, Appli
14	883	55.6	281	US-09-033-337-10	Sequence 10, Appli
15	883	55.6	281	US-09-464-231-10	Sequence 10, Appli
16	256	16.1	451	US-08-287-001A-2	Sequence 2, Appli
17	256	16.1	451	PCT-US95-09941-2	Sequence 2, Appli
18	133.5	8.4	365	US-08-928-383B-23	Sequence 23, Appli
19	133.5	8.4	365	US-08-928-383B-24	Sequence 24, Appli
20	133.5	8.3	387	US-09-175-928-2	Sequence 2, Appli
21	131.5	8.3	365	US-08-928-383B-2	Sequence 2, Appli
22	131.5	8.3	365	US-08-928-383B-26	Sequence 26, Appli
23	131.5	8.3	581	US-08-724-394A-3	Sequence 3, Appli
24	126	7.9	347	US-09-667-135-4	Sequence 4, Appli
25	124.5	7.8	352	US-09-996-243-505	Sequence 505, App
26	124.5	7.8	365	US-08-979-424-3	Sequence 3, Appli
27	124.5	7.8	365	US-09-272-496-2	Sequence 2, Appli

28	120	7.6	581	2	US-08-724-394A-2	Sequence 2, Appli
29	118.5	7.5	398	3	US-09-189-035-6	Sequence 6, Appli
30	118.5	7.5	398	4	US-09-382-086-6	Sequence 6, Appli
31	118.5	7.5	398	4	US-08-999-689A-5	Sequence 5, Appli
32	118	7.4	322	4	US-09-667-135-2	Sequence 2, Appli
33	115.5	7.3	466	4	US-09-604-107A-8	Sequence 8, Appli
34	112.5	7.1	319	3	US-08-597-495B-22	Sequence 22, Appli
35	112.5	7.1	319	3	US-09-068-051A-22	Sequence 22, Appli
36	112.5	7.1	319	4	US-09-336-536-67	Sequence 67, Appli
37	112.5	7.1	319	4	US-09-254-465A-6	Sequence 6, Appli
38	110	6.9	373	4	US-09-926-243-503	Sequence 503, App
39	108	6.8	270	4	US-09-254-465A-24	Sequence 24, Appli
40	108	6.8	273	4	US-09-254-465A-26	Sequence 26, Appli
41	108	6.8	489	4	US-09-667-135-10	Sequence 10, Appli
42	108	6.8	1138	1	US-08-323-474-8	Sequence 8, Appli
43	108	6.8	1138	2	US-08-469-537A-98	Sequence 98, Appli
44	108	6.8	1138	2	US-08-220-240A-5	Sequence 5, Appli
45	105.5	6.6	318	3	US-09-068-051A-32	Sequence 32, Appli

#### ALIGNMENTS

RESULT 1.

US-08-829-525-24

Sequence 24, Application US/08829525

Patent No. 6084083

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/829,525

FILING DATE: 28-MAR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 301 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-829-525-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGQNAVLPCTTPAAPGNLVPVCMGKACPV 60  
DB 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGQNAVLPCTTPAAPGNLVPVCMGKACPV 60  
QY 61 FPGGNVLTERTDERDVNWTSRWYLNQDFKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120  
DB 61 FPGGNVLTERTDERDVNWTSRWYLNQDFKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120  
QY 121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTRGHGPAETQTLGSLPDIINLTQISTLA 180  
DB 121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTRGHGPAETQTLGSLPDIINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATTIRIGIYAGICAGIALALIFGALLIFKWSHSEKKEKIONLSLI 240  
DB 181 NELRDSRLANDLRDSGATTIRIGIYAGICAGIALALIFGALLIFKWSHSEKKEKIONLSLI 240  
QY 241 SLANLPSPGLANAVAGIRSEENITYTIEENVYVEEPEPNYYCYVSSROOPSPLGCRFAM 300  
DB 241 SLANLPSPGLANAVAGIRSEENITYTIEENVYVEEPEPNYYCYVSSROOPSPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

## RESULT 2

US-08-609-583A-24  
Sequence 24, Application US/08609583A  
Patent No. 6204371

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/609,583A

FILING DATE: 01-MAR-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ. ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 301 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-609-583A-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGQNAVLPCTTPAAPGNLVPVCMGKACPV 60  
DB 1 MESHLPFDCVLLLLLLLLLRSSSEVEYRAEVGQNAVLPCTTPAAPGNLVPVCMGKACPV 60  
QY 61 FPGGNVLTERTDERDVNWTSRWYLNQDFKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120  
DB 61 FPGGNVLTERTDERDVNWTSRWYLNQDFKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120  
QY 121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTRGHGPAETQTLGSLPDIINLTQISTLA 180  
DB 121 EKFNKLTVIKPAKVTAPTLQRFDAFPFRLTTRGHGPAETQTLGSLPDIINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSGATTIRIGIYAGICAGIALALIFGALLIFKWSHSEKKEKIONLSLI 240  
DB 181 NELRDSRLANDLRDSGATTIRIGIYAGICAGIALALIFGALLIFKWSHSEKKEKIONLSLI 240  
QY 241 SLANLPSPGLANAVAGIRSEENITYTIEENVYVEEPEPNYYCYVSSROOPSPLGCRFAM 300  
DB 241 SLANLPSPGLANAVAGIRSEENITYTIEENVYVEEPEPNYYCYVSSROOPSPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

## RESULT 3

US-08-937-399-24  
Sequence 24, Application US/08937399  
Patent No. 6288218

## GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/937,399

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

;; INFORMATION FOR SEQ ID NO: 24:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 301 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
US-08-937-399-24

Query Match 100.0%; Score 1587; DB 3; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSHLPDVCVLLLLLLLITRSEVEYRAVGONAYLPCFYTPAAPGNLVPCWGKACPV 60  
DB 1 MFSHLPDVCVLLLLLLLITRSEVEYRAVGONAYLPCFYTPAAPGNLVPCWGKACPV 60

QY 61 FECGNVLTDERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
DB 61 FECGNVLTDERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120

QY 121 EKFNKLVIRPAKTPAPTLQORDTAAPRMLITRGHPARTOTLGSIPDINTLOISTLA 180  
DB 121 EKFNKLVIRPAKTPAPTLQORDTAAPRMLITRGHPARTOTLGSIPDINTLOISTLA 180

QY 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240  
DB 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240

QY 241 SLANLPPSGLANAAVAGIRSEENITYTIEENVYEEBPEYCYVSSRQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAAVAGIRSEENITYTIEENVYEEBPEYCYVSSRQPSQPLGCRFAM 300

QY 301 P 301  
DB 301 P 301

RESULT 4  
US-09-310-367-24  
; Sequence 24, Application US/09310367  
; Patent No. 6414117  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/310.367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829.525  
; FILING DATE: 28-MAR-1997  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A.  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 7853-081  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-790-9090  
;; TELEFAX: 212-869-8864  
;; TELEEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 24:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 301 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
US-09-310-367-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSHLPDVCVLLLLLLLITRSEVEYRAVGONAYLPCFYTPAAPGNLVPCWGKACPV 60  
DB 1 MFSHLPDVCVLLLLLLLITRSEVEYRAVGONAYLPCFYTPAAPGNLVPCWGKACPV 60

QY 61 FECGNVLTDERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
DB 61 FECGNVLTDERDVNWTSRWYLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120

QY 121 EKFNKLVIRPAKTPAPTLQORDTAAPRMLITRGHPARTOTLGSIPDINTLOISTLA 180  
DB 121 EKFNKLVIRPAKTPAPTLQORDTAAPRMLITRGHPARTOTLGSIPDINTLOISTLA 180

QY 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240  
DB 181 NELDSRLANDLRDSGATIRIGIYIGAGICAGLALIFGALIFKWSHSEKIONLSLI 240

QY 241 SLANLPPSGLANAAVAGIRSEENITYTIEENVYEEBPEYCYVSSRQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAAVAGIRSEENITYTIEENVYEEBPEYCYVSSRQPSQPLGCRFAM 300

QY 301 P 301  
DB 301 P 301

RESULT 5  
US-09-032-337-24  
; Sequence 24, Application US/09032337  
; Patent No. 6455685  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032.337  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-032-337-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWGKACPV 60  
DB 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWGKACPV 60  
QY 61 FECSNVLTDRDENVYTSRYWLNQDFRKGVSLEIENVTLADSGIYCCRIQIPIGIMND 120  
DB 61 FECSNVLTDRDENVYTSRYWLNQDFRKGVSLEIENVTLADSGIYCCRIQIPIGIMND 120  
QY 121 EKENLKVIPKAPYTPAPTLQDPFTAAFPRLTTTGHGPAETQTLGSLPDIINTQISTLA 180  
DB 121 EKENLKVIPKAPYTPAPTLQDPFTAAFPRLTTTGHGPAETQTLGSLPDIINTQISTLA 180  
QY 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNLSTLI 240  
DB 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNLSTLI 240  
QY 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 6  
US-09-464-231-24  
Sequence 24, Application US/09464231  
Patent No. 6562343  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/464,231  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-464-231-24

Query Match 100.0%; Score 1587; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 9.3e-155;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWGKACPV 60  
DB 1 MESHLPDVCVLLLLLLLLLRSSEVEYRAEVGNAYLPCFYTPAAPGNLVPCWGKACPV 60  
QY 61 FECSNVLTDRDENVYTSRYWLNQDFRKGVSLEIENVTLADSGIYCCRIQIPIGIMND 120  
DB 61 FECSNVLTDRDENVYTSRYWLNQDFRKGVSLEIENVTLADSGIYCCRIQIPIGIMND 120  
QY 121 EKENLKVIPKAPYTPAPTLQDPFTAAFPRLTTTGHGPAETQTLGSLPDIINTQISTLA 180  
DB 121 EKENLKVIPKAPYTPAPTLQDPFTAAFPRLTTTGHGPAETQTLGSLPDIINTQISTLA 180  
QY 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNLSTLI 240  
DB 181 NEIRDSRLANDLRDSGATTIRIGIYGAGICAGLALALIFGALIFKWSHSEKIKQNLSTLI 240  
QY 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAVAGIRSEENITYTEENVYEEVEEPNEYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 7  
US-08-487-748A-9  
Sequence 9, Application US/08487748A  
Patent No. 5721351  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,748A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-487-748A-9

Query Match 55.6%; Score 883; DB 1; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVEYRAVGQNAVLPCTYTPAAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDEYKVEGKNAYLPCTYTPAAGNLVPCWKGACPV 60  
QY 61 FEGGNVLTREEDVNY-TSRYWNGDFRKGDSLTENTVTLADSGYCCRIQPIGMN 119  
DB 61 SQCTNELRTDERVNTYKSSRYQLKGDNLKGVSLIKNTVLDHGYCCRIQPIGMN 120  
QY 120 DEKNLKIIVAKVTPAPTLQDFTAPFPMLTTRGHGPAETOTIGSLPDLNLTQISTL 179  
DB 121 DKLELKLDIAAKVTPAQTHAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDSGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVGSAGLTLALIIIVLILKWSCKKKLSISL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEENEYCYVSSRQOPS 291  
DB 231 ITLANLPSSGLANAGAVIRSEENIYTIENVEVEEENEYCYVSSRQOPS 281

## RESULT 8

US-08-487-748A-10  
Sequence 10, Application US/08487748A  
Patent No. 5721351  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,748A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-487-748A-10

Query Match 55.6%; Score 883; DB 1; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLTSSSEVEYRAVGQNAVLPCTYTPAAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDEYKVEGKNAYLPCTYTPAAGNLVPCWKGACPV 60  
QY 61 FEGGNVLTREEDVNY-TSRYWNGDFRKGDSLTENTVTLADSGYCCRIQPIGMN 119  
DB 61 SQCTNELRTDERVNTYKSSRYQLKGDNLKGVSLIKNTVLDHGYCCRIQPIGMN 120  
QY 120 DEKNLKIIVAKVTPAPTLQDFTAPFPMLTTRGHGPAETOTIGSLPDLNLTQISTL 179  
DB 121 DKLELKLDIAAKVTPAQTHAGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDSGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIQNLSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVGSAGLTLALIIIVLILKWSCKKKLSISL 230  
QY 240 ISLANLPSSGLANAVABGIRSEENIYTIENVEVEEENEYCYVSSRQOPS 291  
DB 231 ITLANLPSSGLANAGAVIRSEENIYTIENVEVEEENEYCYVSSRQOPS 281

## RESULT 9

US-08-480-070C-10  
Sequence 10, Application US/08480070C  
Patent No. 6066498  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,070C  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-024

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-480-070C-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLRSSSEVEYRAEYVGNAYLPCEYTPAPAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDDYKVEVGNAYLPSCYTLPTSGLTVPWCMKGFCPW 60  
QY 61 FEGGNVLTDERDVNYW-TSRWLNQDFPKGVSLLTENTLADSGIYCCRIQIPGIN 119  
DB 61 SGTNELLTDERNVYOKSSRYOLKGDINKGVSLIKNVTLDDHGYCCRIQIPGLMN 120  
QY 120 DEKENLKVYKPAKTPAPTLQDFPAFPRMLTTRGHGPAETQTLGSLPDINLQISTL 179  
DB 121 DKLEIKLDIKAKTPAOTAHGDSITSPRLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDRSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNTSL 239  
DB 180 ADE-----IKDGETIRTAIHIGVGSAGLTALIGVILKMWSCKKKLSLSL 230  
QY 240 ISLANPPEGLANAVAEGRSEENITYTEENYVEEPEMYCYSSROPS 291  
DB 231 ITLANLPGLANAGAVRIRSEENITYTEENYVEEPEMYCYVNS-CQPS 281

RESULT 10  
US-08-829-525-10  
Sequence 10, Application US/08829525  
Patent No. 6084083  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TITILE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-829-525-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLRSSSEVEYRAEYVGNAYLPCEYTPAPAGNLVPCWKGACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDDYKVEVGNAYLPSCYTLPTSGLTVPWCMKGFCPW 60  
QY 61 FEGGNVLTDERDVNYW-TSRWLNQDFPKGVSLLTENTLADSGIYCCRIQIPGIN 119  
DB 61 SGTNELLTDERNVYOKSSRYOLKGDINKGVSLIKNVTLDDHGYCCRIQIPGLMN 120  
QY 120 DEKENLKVYKPAKTPAPTLQDFPAFPRMLTTRGHGPAETQTLGSLPDINLQISTL 179  
DB 121 DKLEIKLDIKAKTPAOTAHGDSITSPRLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSRLANDRSGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNTSL 239  
DB 180 ADE-----IKDGETIRTAIHIGVGSAGLTALIGVILKMWSCKKKLSLSL 230  
QY 240 ISLANPPEGLANAVAEGRSEENITYTEENYVEEPEMYCYSSROPS 291  
DB 231 ITLANLPGLANAGAVRIRSEENITYTEENYVEEPEMYCYVNS-CQPS 281

RESULT 11  
US-08-609-583A-10  
Sequence 10, Application US/08609583A  
Patent No. 6204371  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TITILE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-609-583A-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFCVLLLLLLLTSSSEVEYRAEYVGNAYLPFCYTPAAGNLYPVCMWGACPV 60  
DB 1 MFSGLTLCVLLLLQLLARSLEDGYKVEGKNAVLPESYTLPTSGTLVPMCMWGACPV 60  
QY 61 FECGNVLRTERDVNYW-TSRWMLNGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
DB 61 SQTNELRTBERNVYQKSSRYQLKGDVSLTIKNTLDDHGTCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPTLQRPDFTAAFPKMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKKELKLDIKAKVTPAQTAAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSLANDLRDSCATIRIGIYGAGICAGLALIFGALLFKWYSHSEKIQNSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVASGLTALILGVLTKWVSCCKKLSLSL 230  
QY 240 ISLANLPPSGLANAVAEGRSEENIYTIENVEVEEPEENEYCYVSSRQPS 291  
DB 231 ITLANLPPGGLANAGAVIRSEENIYTIENVEVEENENYCYVNS-QQPS 281

RESULT 12  
US-08-937-399-10

; Sequence 10, Application US/08937399

; Patent No. 6288218

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; NUMBER OF SEQUENCES: 37 TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/937,399

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/609,583

; FILING DATE: 01-MAR-1996

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-048

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-937-399-10

Query Match 55.6%; Score 883; DB 3; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFCVLLLLLLLTSSSEVEYRAEYVGNAYLPFCYTPAAGNLYPVCMWGACPV 60  
DB 1 MFSGLTLCVLLLLQLLARSLEDGYKVEGKNAVLPESYTLPTSGTLVPMCMWGACPV 60  
QY 61 FECGNVLRTERDVNYW-TSRWMLNGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMN 119  
DB 61 SQTNELRTBERNVYQKSSRYQLKGDVSLTIKNTLDDHGTCCRIQIPGIMN 120  
QY 120 DEKNLKVIRPAKVTAPTLQRPDFTAAFPKMLTTRGHGPAETQTLGSLPDINTQISTL 179  
DB 121 DKKELKLDIKAKVTPAQTAAHGDSTTASPTLTTERNG-SETQTLVTLHNNNGTKISTW 179  
QY 180 ANELRDSLANDLRDSCATIRIGIYGAGICAGLALIFGALLFKWYSHSEKIQNSTL 239  
DB 180 ADE-----IKDSGETIRTAIHIGVASGLTALILGVLTKWVSCCKKLSLSL 230  
QY 240 ISLANLPPSGLANAVAEGRSEENIYTIENVEVEEPEENEYCYVSSRQPS 291  
DB 231 ITLANLPPGGLANAGAVIRSEENIYTIENVEVEENENYCYVNS-QQPS 281

RESULT 13  
US-09-310-367-10

; Sequence 10, Application US/09310367

; Patent No. 6414117

; GENERAL INFORMATION:

; APPLICANT: Levinson, Douglas A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; NUMBER OF SEQUENCES: 38 TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/310,367

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/829,525

; FILING DATE: 28-MAR-1997

; APPLICATION NUMBER: US 08/609,583

; FILING DATE: 01-MAR-1996

; APPLICATION NUMBER: US 08/487,748

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/398,633

; FILING DATE: 03-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7853-081

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-310-367-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAGNLVPCMGKACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCTYLPSTGTVPMCKGKFCPW 60  
QY 61 FPCGNVLTDERDVNYW-TSRVWLNDFPKGDSLTTEENVTLADSGIYCCRIQIPGIN 119  
DB 61 SQTNELLTDERENVYQKSSRYQLKGDINKGDSLLIKNVTLDDHGTCCRIQFGLMN 120  
QY 120 DEKFNKLVIRKAKVTPAFTLORDFTAPFPRMLTTRGHGPAETOTLGSIPDINLTQISTL 179  
DB 121 DKLEELDKIKAKVTPAQTAHGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTISTW 179  
QY 180 ANELDSRLANDLRDGGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
DB 180 ADE-----IKDSEITRTAHIGVGSAGLTALILIGVILIKWSCKKKLSLSL 230  
QY 240 ISLANLPESGLANAVAEGIRSEENITYTTEENVYEEBPNYYCYVSSROOPS 291  
DB 231 ITLANLPESGLANAGAVRIRSEENITYTTEENVYEEVNSNEYCYVNS-CQPS 281

RESULT 14  
US-09-032-337-10  
Sequence 10, Application US/09032337  
Patent No. 6455685  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penite & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,337  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-016

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-032-337-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MFSLPFDVLLLLLLLRSSSEVEYRAEVGNAYLPCFYTPAAGNLVPCMGKACPV 60  
DB 1 MFGSLTNCVLLLLQLLARSLEDGYKVEGKNAVLPCTYLPSTGTVPMCKGKFCPW 60  
QY 61 FPCGNVLTDERDVNYW-TSRVWLNDFPKGDSLTTEENVTLADSGIYCCRIQIPGIN 119  
DB 61 SQTNELLTDERENVYQKSSRYQLKGDINKGDSLLIKNVTLDDHGTCCRIQFGLMN 120  
QY 120 DEKFNKLVIRKAKVTPAFTLORDFTAPFPRMLTTRGHGPAETOTLGSIPDINLTQISTL 179  
DB 121 DKLEELDKIKAKVTPAQTAHGDSTTASPRTLTTERNG-SETQTLVTLHNNNGTISTW 179  
QY 180 ANELDSRLANDLRDGGATIRIGIYGAGICAGLALALIFGALIFKWSHSEKIONLSL 239  
DB 180 ADE-----IKDSEITRTAHIGVGSAGLTALILIGVILIKWSCKKKLSLSL 230  
QY 240 ISLANLPESGLANAVAEGIRSEENITYTTEENVYEEBPNYYCYVSSROOPS 291  
DB 231 ITLANLPESGLANAGAVRIRSEENITYTTEENVYEEVNSNEYCYVNS-CQPS 281

RESULT 15  
US-09-464-231-10  
Sequence 10, Application US/09464231  
Patent No. 6562343  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penite & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/464,231  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048



TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 281 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-09-464-231-10

Query Match 55.6%; Score 883; DB 4; Length 281;  
 Best Local Similarity 63.7%; Pred. No. 1.8e-82;  
 Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

QY 1 MESHLPEDCVLLLLLTSSSEVEYRAEYVGNATLPCTTPAPGNLVPVCKGKACPV 60  
 Db 1 MESHLPEDCVLLLLLTSSSEVEYRAEYVGNATLPCTTPAPGNLVPVCKGKACPV 60  
 QY 61 FECSNVLRTPEDVDYVW-TSRVYMGDPFRKGVSLTENTLADSGIYCCRIQPIGIMN 119  
 Db 61 FECSNVLRTPEDVDYVW-TSRVYMGDPFRKGVSLTENTLADSGIYCCRIQPIGIMN 119  
 QY 120 DEKFNKLVIRPAKVPAPTLQRODFAFPMLTTRGHGPAETQTLSLPDINLTQISTL 179  
 Db 120 DEKFNKLVIRPAKVPAPTLQRODFAFPMLTTRGHGPAETQTLSLPDINLTQISTL 179  
 QY 121 DKKLELKDIKAQKTPAQTAHGDSTTASPRITLTERNG-SETOTLVTLHNNNGIKISTW 179  
 Db 121 DKKLELKDIKAQKTPAQTAHGDSTTASPRITLTERNG-SETOTLVTLHNNNGIKISTW 179  
 QY 180 ANELRDSRLANDLSDGATIRIGIYTGAGICGALALIFGALIFKWTSHSKEKIQNLSTL 239  
 Db 180 ANELRDSRLANDLSDGATIRIGIYTGAGICGALALIFGALIFKWTSHSKEKIQNLSTL 239  
 QY 240 ISLANLPSPGLANAVAEGRSEENITYTEENYVEEENPYCYVSSROOPS 291  
 Db 240 ISLANLPSPGLANAVAEGRSEENITYTEENYVEEENPYCYVSSROOPS 291  
 QY 231 ITLANLPSPGLANAGAVRIRSEENITYTEENYVEEENPYCYVSSROOPS 281  
 Db 231 ITLANLPSPGLANAGAVRIRSEENITYTEENYVEEENPYCYVSSROOPS 281

Search completed: November 22, 2003, 05:45:21  
 Job time : 21.6529 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:43:31 ; Search time 44.9948 Seconds  
(without alignments)  
1221.260 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MFSHLPFDCVLLLLLLTR.....CYVSRQPSQLGCRFAMP 301

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*\n2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*\n3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*\n4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*\n5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*\n6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*\n7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*\n8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*\n9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*\n10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*\n11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*\n12: /cgn2\_6/ptodata/1/pubpaa/US09D\_NEW\_PUB.pep:\*\n13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*\n14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*\n15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*\n16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*\n17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*\n18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	301	US-10-252-131-24	Sequence 24, Appl
2	1587	100.0	301	US-10-004-633-24	Sequence 24, Appl
3	1587	100.0	301	US-10-188-012-31	Sequence 31, Appl
4	1581	99.6	301	US-10-188-012-39	Sequence 29, Appl
5	940	59.2	185	US-09-529-063-34	Sequence 34, Appl
6	940	59.2	185	US-10-414-378-34	Sequence 34, Appl
7	885	55.8	281	US-10-188-012-11	Sequence 11, Appl
8	883	55.6	281	US-10-252-131-10	Sequence 10, Appl
9	883	55.6	281	US-10-004-633-10	Sequence 9, Appl
10	883	55.6	281	US-10-188-012-9	Sequence 9, Appl
11	718	45.2	142	US-09-966-546-22	Sequence 22, Appl
12	718	45.2	142	US-09-966-545-22	Sequence 22, Appl
13	718	45.2	142	US-09-965-212-22	Sequence 22, Appl
14	718	45.2	142	US-10-189-940-22	Sequence 22, Appl
15	718	45.2	142	US-10-189-940-146	Sequence 146, App

16	613	38.6	125	16	US-10-189-940-145	Sequence 145, App
17	333	21.0	282	15	US-10-188-012-3	Sequence 3, Appl
18	323.5	20.4	305	15	US-10-188-012-1	Sequence 1, Appl
19	322	20.3	359	15	US-10-188-012-17	Sequence 17, Appl
20	322	20.3	359	15	US-10-188-012-19	Sequence 19, Appl
21	322	20.3	359	15	US-10-188-012-23	Sequence 23, Appl
22	319.5	20.1	364	15	US-10-188-012-25	Sequence 25, Appl
23	319	20.1	365	15	US-10-188-012-21	Sequence 21, Appl
24	317.5	20.0	364	15	US-10-188-012-27	Sequence 27, Appl
25	273.5	17.2	305	15	US-10-188-012-5	Sequence 5, Appl
26	273.5	17.2	305	15	US-10-188-012-7	Sequence 7, Appl
27	250.5	15.8	345	15	US-10-188-012-13	Sequence 13, Appl
28	246.5	15.5	345	15	US-10-188-012-15	Sequence 15, Appl
29	245.5	15.5	378	15	US-10-188-012-35	Sequence 35, Appl
30	245.5	15.5	378	15	US-10-188-012-35	Sequence 35, Appl
31	245.5	15.5	379	11	US-09-813-153-138	Sequence 138, App
32	238.5	15.0	183	9	US-09-739-907-65	Sequence 65, App
33	132.5	8.3	387	14	US-10-114-893-133	Sequence 133, App
34	132.5	8.3	387	15	US-10-016-249-2	Sequence 2, Appl
35	131.5	8.3	523	10	US-09-910-174A-11	Sequence 11, Appl
36	131.5	8.3	523	10	US-09-855-866-10	Sequence 10, Appl
37	131.5	8.3	523	10	US-09-896-738-16	Sequence 16, Appl
38	131.5	8.3	523	12	US-10-087-887-53	Sequence 53, Appl
39	127.5	8.0	365	9	US-09-899-634A-4	Sequence 4, Appl
40	126	7.9	347	15	US-10-281-478-9	Sequence 9, Appl
41	125	7.9	505	12	US-10-114-153-12	Sequence 12, Appl
42	124.5	7.8	261	9	US-09-889-634A-2	Sequence 2, Appl
43	124.5	7.8	352	9	US-09-989-722-505	Sequence 505, App
44	124.5	7.8	352	9	US-09-989-723-505	Sequence 505, App
45	124.5	7.8	352	9	US-09-989-279-505	Sequence 505, App

## ALIGNMENTS

RESULT 1  
US-10-252-131-24

Sequence 24, Application US/10252131

Publication No US20030158399A1

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/252,131

FILING DATE: 20-Sep-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Cornuzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-10-252-131-24

Query Match 100.0%; Score 1587; DB 12; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
QY 241 SLANLPPSGLANAVAEGIRSEENIYTIENVEVEEPNEYYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAVAEGIRSEENIYTIENVEVEEPNEYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 2  
US-10-004-633-24

Sequence 24, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 301  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-004-633-24

Query Match 100.0%; Score 1587; DB 15; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
QY 241 SLANLPPSGLANAVAEGIRSEENIYTIENVEVEEPNEYYCYVSSROQPSQPLGCRFAM 300  
DB 241 SLANLPPSGLANAVAEGIRSEENIYTIENVEVEEPNEYYCYVSSROQPSQPLGCRFAM 300  
QY 301 P 301  
DB 301 P 301

RESULT 3  
US-10-188-012-31

Sequence 31, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Dekruyf, Rosemarie  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 31  
LENGTH: 301  
TYPE: PRT  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(301)  
OTHER INFORMATION: TIM-3, allele 2  
US-10-188-012-31

Query Match 100.0%; Score 1587; DB 15; Length 301;  
Best Local Similarity 100.0%; Pred. No. 2.7e-151;  
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
DB 1 MFSLPFDVCVLLLLLLLTSSEVEYRAEVGONAYLPCFYTTPAAGNLVPCWGKACPV 60  
QY 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
DB 61 FECGNVVLRTDERDVNWTSRWYLNQDFRKGVSLTIENVTLADSGIYCCRIQIGIMND 120  
QY 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
DB 121 EKFNKLVIKPAKVTAPTLQRDFTAAFPRLMTTRGHGPAETQTGSLPDINLTQISTLA 180  
QY 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240  
DB 181 NELRDSRLANDLRDSCGATIRIGIYGAGICAGLALALIFGALIFKMYSHSKEKIQNLSLI 240

Db 181 NELSDSRANDLRDSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
QY 241 SLANLPPSGLANAAVEGIRSEENIYTIENENYEEVEENEEYCYSSROQPSQPGCRPFAM 300  
Db 241 SLANLPPSGLANAAVEGIRSEENIYTIENENYEEVEENEEYCYSSROQPSQPGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 4  
US-10-188-012-29  
; Sequence 29, Application US/10188012  
; Publication No. US20030124114A1  
; GENERAL INFORMATION:  
; APPLICANT: McInliffe, Jennifer Jones  
; APPLICANT: Umetsu, Dale T.  
; APPLICANT: Dekruyff, Rosemarie  
; APPLICANT: Kuchroo, Vijay  
; APPLICANT: Freeman, Gordon J.  
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
; FILE REFERENCE: STAN-235  
; CURRENT APPLICATION NUMBER: US/10/188,012  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: TIM-3, allele 1  
US-10-188-012-29

Query Match 99.6%; Score 1581; DB 15; Length 301;  
Best Local Similarity 99.7%; Pred. No. 1.1e-150;  
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60  
QY 61 FECGNVLRTERDENVWYTSRYWLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECGNVLRTERDENVWYTSRYWLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQISTLA 180  
Db 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQISTLA 180  
QY 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQISTLA 180  
Db 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQISTLA 180  
QY 181 NELSDSRANDLRDSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
Db 181 NELSDSRANDLRDSGATIRIGIYIGAGICAGLALALIFGALIPKWSHSEKIQNSLI 240  
QY 241 SLANLPPSGLANAAVEGIRSEENIYTIENENYEEVEENEEYCYSSROQPSQPGCRPFAM 300  
Db 241 SLANLPPSGLANAAVEGIRSEENIYTIENENYEEVEENEEYCYSSROQPSQPGCRPFAM 300  
QY 301 P 301  
Db 301 P 301

RESULT 5  
US-09-529-063-34  
; Sequence 34, Application US/09529063

; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAIKICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-063-34

Query Match 59.2%; Score 940; DB 10; Length 185;  
Best Local Similarity 99.4%; Pred. No. 1.3e-86;  
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60  
QY 61 FECGNVLRTERDENVWYTSRYWLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
Db 61 FECGNVLRTERDENVWYTSRYWLNQDPRKGDVSLTIENVTLADSGIYCCRIQIPGIMND 120  
QY 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQI 176  
Db 121 EKFNKLVIKPAKTPAPFTLQDFTAPFPMULTTRGHGPAETQTLGSLPDIINTLQI 176

RESULT 6  
US-10-414-378-34  
; Sequence 34, Application US/10414378  
; Publication No. US20030165981A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAIKICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/10/414,378  
; PRIOR APPLICATION NUMBER: US/09/529,063  
; PRIOR FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-414-378-34

Query Match 59.2%; Score 940; DB 12; Length 185;  
Best Local Similarity 99.4%; Pred. No. 1.9e-86;  
Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60  
Db 1 MFSLPPDCVLLLLLLLITRSSEVEYRAEVGQNNAYLPCFTTPAAPGNLVPVCMKGACPV 60

Qy	Db	Qy	Db
61	61	121	121
EEGNNVLTDRDVAWYMTSRXYLANDPFRGVDVSLTENTVLTADSGIYCCRIQIPGMND 120	EEGNNVLTDRDVAWYMTSRXYLANDPFRGVDVSLTENTVLTADSGIYCCRIQIPGMND 120	EEFNKLKLYKPAKVPAPFLTORDFTAAPRMLTTRHGPAETQTLGSLPINITOI 176	EEFNKLKLYKPAKVPAPFLTORDFTAAPRMLTTRHGPAETQTLGSLPINITOI 176

```

RESULT 7
US-10-188-012-11
; Sequence 11, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetani, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchnroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 281
; TYPE: prt
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(281)
; OTHER INFORMATION: TIM-3, C.D2 ES-HBA and DBA/2J allele
US-10-188-012-11

```

Query Match	55.8%	Score	885	DB	15	Length	281
Best Local Similarity	64.0%	Pred. No.	1,2e-80				
Matches	187	Conservative	26	Mismatches	67	Indels	12
				Gaps	4		
QY	1	MFSHLPDVCVLLMLLLLTTRSSSEVRYRAVEGONAVLPCTYPAPAGNVLYVCWGKACPV	60				
Db	1	MFSGLTNCVLLLTLLRLARSLENAVVEFGNAALPCSYTLSTGVALPVMCMGKFCPM	60				
QY	61	FECCGVVLRTERDENVVM-TSRVMNGDPRKCDVSLTEENTVLADSGICCRIGIPIGNN	119				
Db	61	SGCTNELLRTBERNVTYQSSRYOLKGLDKNKDVSLLIKNVTLLDHDGYCCGIIQPGMLN	120				
QY	120	DEKFLTKLYIKAAKTPAPLTORDTPAAFPKMLTTRHGCPATOTLGSLPIDNLQISTL	179				
Db	121	DKKELTKLDIKAAKTPAOTAHAGSDTSPRLTLTERNG-STQTLVTLLHNNNGKRISTW	179				
QY	180	ANELDSRLANDIRDGATIRIGIYIGAGICAGLALALFAGALIFKMTSHSKEKTONTSL	239				
Db	180	ADE-----IKDSGETIRPAIHIGVCSAGLTLLALITGLLILKMTSCKKKUSLSL	230				
QY	240	ISLANLPBSGLANAAVEGIRSEENITYTEENRYVEBEPEHYECYVSSRQPS	231				
Db	231	ITLANLPBSGLANAAVIRISEENITYTEENRYEVENSYECYVNS-QQPS	281				

RESULT 8  
 US-10-252-131-10  
 : Sequence 10, Application US/10252131  
 : Publication No. US20030158399A1  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Levinson, Douglas A.  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 : TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 :  
 : NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-0909  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-252-131-10

Query Match	55.6%	Score 883	DB 12	Length 281
Best Local Similarity	63.7%	Pred. No. 1,9e+80		
Matches 186	Conservative 26	Mismatches 68	Indels 12	Gaps 4
QY	1	MFSLPDCVTLILLILLITRSSEVYRAEVGONATLPCTYPAPAGNLYPVCWGKACPV	60	
Db	1	MFSGTLTNCVLLLTQTLARSLDEQKYKVEGNATLPCTYTLPTSGTLVPMCKGKFCPW	60	
QY	61	FECCGVAVRTDERDVNYW-TSRWYMGDPFKKDVSLTTEENTVLADSGICCRITQFPGINN	119	
Db	61	SGCTNELLRTERBNVYQKSRSTYQKGLNKGDSVLLIKNVLDDHGTCCRIQFPGINN	120	
QY	120	DEKFLKLVYPAKTPAPTLORDPTAAFPRLMTTRGHPAORTQIGSLPDINTLOISTL	179	
Db	121	DKKELTKDIDIAAKYTPAQTAHGDSTYASPRLLTTERBG-SETQTLVYLTANNNGTKISW	179	
QY	180	ANELDSRLANDRDSGATIRIGIYIGAGICAGLALALFQALIFKWTSHSKKIONLSL	239	
Db	180	ADE-----IKSGETIRAHNIIGVAGSAGLTALILIGVLILKWTSCKKKXSSLSL	230	
QY	240	ISLANLPFSGLANNAVEGIRSEENITYTTEENVYEVPEPEHYCYVSSROOPS	291	
Db	231	ITLANLPFGGLANNAVARIISEENITYTTEENVYEVENSSEYCYVNS-QOPS	281	

RESULT 9  
US-10-004-633-10  
Sequence 10, Application US/10004633-10  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
FILE OF INVENTION: DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-125  
CURRENT APPLICATION NUMBER: US/10/004,633  
PRIOR APPLICATION NUMBER: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 10  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-004-633-10

Query Match 55.6%; Score 883; DB 15; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.9e-80;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAEVGQNAVLPFCYTPAAPAGNLVPCWKGACPV 60  
1 MFSGLTNCVLLLLQLLARSLLEDGYKVEGKNALPSCSYTLPTSGTLVPCWKGACFPW 60  
Db 61 FECGNVLTARDVDVNTWTSRYWLNDFRKGDSLTTEVNTLADSGIYCCRIQIPGIMN 119  
61 SQCTNELRTDERNVTYQKSSRYQKDLNKGVSLLIKNTLDDHGTGCCRIQIPGIMN 120  
Qy 120 DEKNLKVIRPAKVTAPTLQDRFTAAPPMLTTRGHPAETQTLGSLPINTLOISTL 179  
121 DKKELEKLDIAAKVTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
Db 180 ANELDRSLANDLDSGATIRIGIYGAGICAGLALIFGALLFKWYSHSKKXKIQNL 239  
180 ADE-----IKDSGTIRTAHIGVGSAGLTLALLIGVLLKWSCKKKLSLSL 230  
Qy 240 ISLANLPSSGLANAVASGIRSEENIYITEENVYEEVENEYCYVSSRQOPS 291  
231 ITLANLPSSGLANAGAVRIRSEENIYITEENVYEEVENEYCYVNS-QOPS 281  
Db

RESULT 10  
US-10-188-012-9  
Sequence 9, Application US/10188012  
Publication No. US20030124114A1  
GENERAL INFORMATION:  
APPLICANT: McIntire, Jennifer Jones  
APPLICANT: Umetsu, Dale T.  
APPLICANT: Kuchroo, Vijay  
APPLICANT: Freeman, Gordon J.  
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
FILE REFERENCE: STAN-235  
CURRENT APPLICATION NUMBER: US/10/188,012  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(281)  
OTHER INFORMATION: TIM-3 BALB/c allele

US-10-188-012-9

Query Match 55.6%; Score 883; DB 15; Length 281;  
Best Local Similarity 63.7%; Pred. No. 1.9e-80;  
Matches 186; Conservative 26; Mismatches 68; Indels 12; Gaps 4;

Qy 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAEVGQNAVLPFCYTPAAPAGNLVPCWKGACPV 60  
1 MFSGLTNCVLLLLQLLARSLLEDGYKVEGKNALPSCSYTLPTSGTLVPCWKGACFPW 60  
Db 61 FECGNVLTARDVDVNTWTSRYWLNDFRKGDSLTTEVNTLADSGIYCCRIQIPGIMN 119  
61 SQCTNELRTDERNVTYQKSSRYQKDLNKGVSLLIKNTLDDHGTGCCRIQIPGIMN 120  
Qy 120 DEKNLKVIRPAKVTAPTLQDRFTAAPPMLTTRGHPAETQTLGSLPINTLOISTL 179  
121 DKKELEKLDIAAKVTPAQTAGDSTTASPTLTTERNG-SETOTLVTLHNNNGTKISTW 179  
Db 180 ANELDRSLANDLDSGATIRIGIYGAGICAGLALIFGALLFKWYSHSKKXKIQNL 239  
180 ADE-----IKDSGTIRTAHIGVGSAGLTLALLIGVLLKWSCKKKLSLSL 230  
Qy 240 ISLANLPSSGLANAVASGIRSEENIYITEENVYEEVENEYCYVSSRQOPS 291  
231 ITLANLPSSGLANAGAVRIRSEENIYITEENVYEEVENEYCYVNS-QOPS 281  
Db

RESULT 11

US-09-966-546-22  
Sequence 22, Application US/09966546  
Patent No. US20020168716A1  
GENERAL INFORMATION:

APPLICANT: Fernandes, Elma  
APPLICANT: Shinkens, Richard A.  
TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encoding  
FILE REFERENCE: Cura-46 (15966-546)  
CURRENT APPLICATION NUMBER: US/09/966,546  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 09/544,511  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 22  
LENGTH: 142  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-966-546-22

Query Match 45.2%; Score 718; DB 10; Length 142;  
Best Local Similarity 97.8%; Pred. No. 3e-64;  
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MFSHLPDCCVLLLLLLLTSSSEVEYRAEVGQNAVLPFCYTPAAPAGNLVPCWKGACPV 60  
1 MFSGLTNCVLLLLQLLARSLLEDGYKVEGKNALPSCSYTLPTSGTLVPCWKGACFPW 60  
Db 61 FECGNVLTARDVDVNTWTSRYWLNDFRKGDSLTTEVNTLADSGIYCCRIQIPGIMN 120  
61 SQCTNELRTDERNVTYQKSSRYQKDLNKGVSLLIKNTLDDHGTGCCRIQIPGIMN 120  
Qy 121 EKFNKLVIRPAKVT 135  
121 EKFNKLVIRKPEWT 135  
Db

RESULT 12

US-09-966-545-22  
Sequence 22, Application US/09966545  
Patent No. US20020172999A1  
GENERAL INFORMATION:  
APPLICANT: Fernandes, Elma

```
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20020172999A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-545-22
```

```
Query Match      45.2%; Score 718; DB 10; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60
        |||||
Db       1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60

QY      61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||||
Db       61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120

QY      121 EKFNKLVIKPAKVT 135
        |||||
Db       121 EKFNKLVIKRGEMT 135
```

```
RESULT 13
US-09-965-212-22
; Sequence 22, Application US/09965212
; Publication No. US20030003462A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20030003462A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USSN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-22
```

```
Query Match      45.2%; Score 718; DB 11; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60
        |||||
Db       1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60

QY      61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||||
Db       61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120

QY      121 EKFNKLVIKPAKVT 135
        |||||
Db       121 EKFNKLVIKPAKVT 135
```

```
Db       121 EKFNKLVIKRGEMT 135
```

```
RESULT 14
US-10-189-940-22
; Sequence 22, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
; APPLICANT: Anderson, David
; APPLICANT: Padigaru, Muraidhara
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
; APPLICANT: Casman, Stacie
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: No. US20030129613A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 22
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-22
```

```
Query Match      45.2%; Score 718; DB 16; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60
        |||||
Db       1 MFSLPFDVLLLLLLTRSSSEVEYRAEVGQNAVLPCEFTPAAPGNLVPCWGKACPV 60

QY      61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
        |||||
Db       61 FECGNVVLRTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120

QY      121 EKFNKLVIKPAKVT 135
        |||||
Db       121 EKFNKLVIKRGEMT 135

RESULT 15
US-10-189-940-146
; Sequence 146, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
```



```

; APPLICANT: Anderson, David
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
; APPLICANT: Caeman, Stacie
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: No. US20030129613A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: Cuiaseqlist version 0.1
; SEQ ID NO 146
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-146

Query Match          45.2%; Score 718; DB 16; Length 142;
Best Local Similarity 97.8%; Pred. No. 3e-64;
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSLPFDVCVLLILLILTRSEVEYRAVGONAYLPCTTPAPAGNLVPCWKGACPV 60
Db 1 MFSLPFDVCVLLILLILTRSEVEYRAVGONAYLPCTTPAPAGNLVPCWKGACPV 60

QY 61 FECGNVVLRTDPRDYNWTSRYWLNGLDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120
Db 61 FECGNVVLRTDPRDYNWTSRYWLNGLDFRKGDVSLTIENVTLADSGIYCCRIQIPIGIMND 120

QY 121 EKFNKLVIKPAKVT 135
Db 121 EKFNKLVIKPGWT 135

```

Search completed: November 22, 2003, 05:56:16  
 Job time : 46.9948 secs

**THIS PAGE BLANK (USPTO)**

DE Similarity hypothetical protein FLJ14428.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Euteheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=lung;  
RA Straube R.;  
RL Submitted (Jan-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC020843; AAH20843.1; -  
DR InterPro; IPR003599; IG-  
DR InterPro; IPR007110; IG-like.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 142 AA; 16149 MW; E2C82D8B5AABC23E CRC64;

Query Match 45.2%; Score 718; DB 4; Length 142;  
Beet Local Similarity 97.8%; Pred. No. 4.9e-61;  
Matches 132; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSLPFCVLLLLLLLTSSSEVEYRAVGONAYLPFPTPAAPAGNLVPCMGKACPV 60  
Db 1 MFSHLPFCVLLLLLLLTSSSEVEYRAVGONAYLPFPTPAAPAGNLVPCMGKACPV 60  
QY 61 FEGGNVVLRTDERDVNYVTSRWINGDERKDVSLTIENVTLLADSGIYCCRIQPIGIND 120  
Db 61 FEGGNVVLRTDERDVNYVTSRWINGDERKDVSLTIENVTLLADSGIYCCRIQPIGIND 120  
QY 121 EKFNLLKLVYRKPKYT 135  
Db 121 EKFNLLKLVYRKPKYT 135  
QY 121 EKFNLLKLVYRKPKYT 135  
Db 121 EKFNLLKLVYRKPKYT 135

RESULT 5  
054947 PRELIMINARY; PRT; 307 AA.  
AC 054947;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Kidney injury molecule-1 precursor (KIM-1).  
GN KIM-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;  
RX MEDLINE=98129827; PubMed=9461608;  
RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,  
RA Cate R.L., Sanicola M.;  
RT "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion  
RT molecule containing a novel immunoglobulin domain, is up-regulated in  
RT renal cells after injury.";  
RL J. Biol. Chem. 273:4135-4142(1998).  
CC -1- FUNCTION: MAY ACT AS AN EPITHELIAL CELL ADHESION MOLECULE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN LIVER, SPLEEN AND  
CC NORMAL KIDNEY. LEVELS INCREASE IN THE POSTISCHEMIC KIDNEY WITH  
CC EXPRESSION FOUND IN REGENERATING PROXIMAL TUBULE EPITHELIAL CELLS.  
CC -1- INDUCTION: IN RENAL CELLS, AFTER INJURY.  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC ONE V-LIKE DOMAIN.  
DR EMBL; AF035963; AAC3546.1; -  
DR InterPro; IPR003599; IG-  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003066; IG\_MHC.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.

KM	Transmembrane; Cell adhesion; Glycoprotein; Signal.
FT	SIGNAL
FT	CHAIN
FT	CHAIN
FT	DOMAIN
FT	TRANSMEM
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
SO	SEQUENCE
QY	Query Match
D8	Best Local Similarity 32.5%; Pred. No. 8e-24;
MATCHES	96; Conservative 43; Mismatches 116; Indels 40; Gaps 10;
D8	11 LLLLLLTTSSSEVEVAEAGNAYLPCFTTPAARGLVVCGKGCAPVEGNNVLRT 70
QY	71 DERDVNWTS-RWLANDPFRKDDSLTLENTVLADSGIYCRRIGIPIMDEKNLKVI 129
D8	68 NGVYVTRSSGRNVIKIRISGEVDLSLTENSVDSDSLCYCRVAIVPFMFNDOKTFSLIEV 127
QY	130 KPAKVTPAPTLPDRDF--TAAPPRMLLTRG-HGAPEQTGLSPDINTLQ-----ISTL-A 180
D8	128 KPEIPTSPPTRPPTTRPTTRPTTSTRSHVPSTRAVSTSTPTEGTQTHKPEITTFYA 187
QY	181 NEL-----RDSRLANDLRDSGATRI-----GIYGAGICACLA 214
D8	188 HETTAETETPSYPADPMNGTIVSSEBAMNHVTRIPLRKRQRNPRTGFYGMGV-AALL 246
QY	215 LATIFGALIFKWYSHKSKEKIONLSILSIANLPPGGLANAVALAEGRISENTYITE 269
D8	247 LLHLASTVVTRYIIIRKKMGSLFVAFPHYSKSKALONAAIVHRAEDNIYTIED 301

RESULT 6

ID	O8VIM1	PRELIMINARY;	PRT:	282 AA.
AC	O8VIM1;			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	TIM1.			
GN	TIM1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
R2	SEQUENCE FROM N.A.			
RC	STRAIN=DBA/2; TISSUE=Spleen;			
RA	McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,			
RT	"Tadp, a major T cell regulatory locus that controls the development			
RT	of airway hyperreactivity, cosegregates with variants in a novel gene			
RT	family.";			
KL	Submitted (JUL-2001) to the EMBL/GenBank/DDJB databases.			
DR	EMBL; AF399830; AAL35775.1; -			
DR	MCD; MG1.2159680; Timd1.			
DR	InterPro; IPR003599; IG.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR003006; IG_MHC.			
DR	Pfam; PF00047; IG_1.			
DR	SMART; SMO0409; IG_1.			

DR PROSITE: PSS0835; IG\_LIKE; 1.  
SQ SEQUENCE 282 AA; 30968 MM; 7D30E0698F0AC5F CRC64;  
Query Match 21.0%; Score 333; DB 11; Length 282;  
Best Local Similarity 33.8%; Pred. No. 1.4e-23;  
Matches 93; Conservative 36; Mismatches 124; Indels 22; Gaps 8;  
QY 12 LLLLLLLTSSEVEYAEVGNAYLPCFYTPAAGNLVPCMGKACPVFEGCNVLRD 71  
DB 11 LLLLEAGAVSYEVEGVGVHPVTLPCYSTR--GITTTCMGRCGQCPSSACNTLIMTN 68  
QY 72 ERDVNYM-TSRWYLNAGDFRKGVSLTIENVTLADSGIYCCRIQIPGINDEKFNLLVIK 130  
DB 69 GHRVTVQKSSRYLKHISBGVSLTIENSVSDSGLYCCRAVEIPGENDQKTFSLQYK 128  
QY 131 ---PAKYTPAPTLQORDFTAFAPRMLTTRG-HGPAETQLGSLPDIMLTQISTLANELRDS 186  
DB 129 PEIPTRPPRPRTTRTATGRPTTISTRSTHVTSRVSSTPPTS-THTWTHKPDWNGT 187  
QY 187 RLANDLRDSGATRI-----GIYAGIC-AGLALIFGALLFKWYSHSKKXI 234  
DB 188 VTSSGDTWSNHTAIPPGKPKQNPTRKGFYV--GICIAALLLLVSTVAITRYILMKRKS 245  
QY 235 ONLSLISLANLPSSGLANVAEGIRSEENIYITEE 269  
DB 246 ASLSVAVFRVSKIEALONAAVHSHRADDNIYIED 280  
RESULT 7  
ID Q8VIM2 PRELIMINARY; PRT; 305 AA.  
AC Q8VIM2  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE TIM1.  
GN TIMD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BA16/c; TISSUE=Spleen;  
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,  
RA Freeman G.J., Umetsu D.T., Detryiff R.H.;  
RT "Tap1, a major T cell regulatory locus that controls the development  
RT of airway hyperactivity, cosegregates with variants in a novel gene  
RT family".  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF399829; AL35774.1; -.  
DR MGD; MGI:2159680; Tmd1.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; Ig; 1.  
DR PROSITE; PSS0835; IG\_LIKE; 1.  
SQ SEQUENCE 305 AA; 33391 MM; 8F4EA38627FE85FB CRC64;  
Query Match 20.4%; Score 323.5; DB 11; Length 305;  
Best Local Similarity 31.0%; Pred. No. 1.3e-22;  
Matches 93; Conservative 35; Mismatches 123; Indels 49; Gaps 9;  
QY 12 LLLLLLLTSSEVEYAEVGNAYLPCFYTPAAGNLVPCMGKACPVFEGCNVLRD 71  
DB 11 LLLLEAGAVSYEVEGVGVHPVTLPCYSTR--GITTTCMGRCGQCPSSACNTLIMTN 68  
QY 72 ERDVNYM-TSRWYLNAGDFRKGVSLTIENVTLADSGIYCCRIQIPGINDEKFNLLVIK 130  
DB 69 GHRVTVQKSSRYLKHISBGVSLTIENSVSDSGLYCCRAVEIPGENDQKTFSLQYK 128  
QY 131 ---PAKYTPAPTLQORDFTAFAPRMLTTRG-HGPAETQLGSLPDIMLTQISTLANELRDS 186

DB 129 PEIPTRPPRPRTTRTATGRPTTISTRSTHVPSIRVSTSTP---TSTHTWKPEPT 185  
QY 187 RLA-----NDLRDSGAT-----IRIGIYAGIC 210  
DB 186 TCEPHTTAETVATGIPSHPTDWNQGVTSGGDTWSNHTAIPPGKPKQNPTRKGFYV--GIC 243  
QY 211 -AGLALIFGALLFKWYSHSKKXIQLNLSLISLANLPSSGLANVAEGIRSEENIYITEE 269  
DB 244 IALLLLLVSTVAITRYILMKRKSLSVAVFRVSKIEALONAAVHSHRADDNIYIED 303  
RESULT 8  
ID Q43656 PRELIMINARY; PRT; 359 AA.  
AC Q43656  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hepatitis A virus cellular receptor 1.  
GN HAVCR-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA MEDLINE=98325180; PubMed=9658108;  
RA Feigelsstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;  
RT "The human homolog of HAVCR-1 codes for a hepatitis A virus cellular  
RT receptor".  
RL J. Virol. 72:6621-6628 (1998).  
DR EMBL; AF043724; AAC39862.1; -.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; Ig; 1.  
DR PROSITE; PSS0835; IG\_LIKE; 1.  
DR PROSITE; PSS0290; IG\_MHC; 1.  
KM Receptor.  
SQ SEQUENCE 359 AA; 38704 MM; C207FEC562DC62CA CRC64;  
Query Match 20.3%; Score 322; DB 4; Length 359;  
Best Local Similarity 27.9%; Pred. No. 2.2e-22;  
Matches 102; Conservative 43; Mismatches 110; Indels 110; Gaps 13;  
QY 10 VLLLLLLLTRS--SEVEYAEVGNAYLPCFYTPAAGNLVPCMGKACPVFEGCNV 67  
DB 6 VLSLTLHLADSVAGSVKVGEGSPVTLPCYIS---GAVTSMCNRRSCSLFTQONCI 61  
QY 68 LRTDERDVNY-WTSRYWLNAGDFRKGVSLTIENVTLADSGIYCCRIQIPGINDEKFNLLK 126  
DB 62 WVTNGHVTYRKDTYKLLGDLSSRDVSLTIENAVSDSGYCCVEHKGWENDMKITVS 121  
QY 127 LVIKPAKVPAP-----TLQORDFTAFAPRMLTTR-- 154  
DB 122 LEIVPCKVTTTIVTVPTVTRSTVPTTIVPTTVPTTIVPTTIVPTTIVPTTIVPTT 181  
QY 155 -----RHGPAETQLGSLPDIMLTQISTLAN 181  
DB 182 TTSVPTTISIPPTTSVPVTTVSTVPPMPRLPRONHEVAISP--SSPPAETHPTTLOG 239  
QY 182 ELR-----DSRLANDLRDSGAT-----RIGIYAGIC 211  
DB 240 AIRREBTSPLSYTTDGDVTESSDGLMNNNQTLPLEHSLTANNTKGIY--AGVCI 297  
QY 212 G-LALALIFGALLFKWYSHSKKXIQLNLSLISLANLPSSGLANVAEGIRSEENIYITEE 270  
DB 298 SYLVIALALGVIAIKYFPFKE--VOOLS--VSFSLQIKALQNAVEREQADENIY-IENS 354  
QY 271 VVEVE 275

Db 355 LYATD 359

## RESULT 9

Q96D42 PRELIMINARY; PRT; 364 AA.  
 AC Q96D42; 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013325; AAH13325.1; -.  
 DR Genew; HGNC:17866; HAVER1.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 364 AA; 39249 MW; AA67C7DCTFAC91F1 CRC64;

Query Match 19.9%; Score 316.5; DB 4; Length 364;  
 Best Local Similarity 27.6%; Pred. No. 7.6e-22;  
 Matches 102; Conservative 42; Mismatches 111; Indels 115; Gaps 12;

QY 10 VLLLLLLLTSS--SEVERAEVGNAYLPCFYTPAAGNLPVPCMGKACVFECCGNV 67  
 DB 6 VILSLILHLADSVAGSVKVGEGAGSPVLPCHYS---GAVTSMKWNKSGSLFTCCQNGI 61  
 QY 68 LRTDERDVNY-WTSRYWNGDFPKGDVSLTIENTVTLADSGIYCCRIQIPGIMNDEKFNK 126  
 DB 62 VWTNGTHTVYKRDTRYKLLGDLSDRDVSLTIENTAVDSGVYCCVEHKGWENDKITVS 121  
 QY 127 LVIRAKATTPAFTLORDTAAFPRLTT----- 154  
 DB 122 LEIVPCKVTTTPVITVPTVTRSTTVPTTTPVMTTVPPTTTPPTTMSIPTTTVLTT 181  
 QY 155 -----RGHGAETQTLGSLPDINLQI 176  
 DB 182 MVTSTTTSVPTTTSIPTTTSVPTTTSVTPVPMPLPRONHEPVATSP--SSPOPAETHP 239  
 QY 177 STLANELR---DSRLANDLRDSGAT-----RIGIYIG 206  
 DB 240 TTLQGAIRREPTSSPLSYTTDGNVTVESSDGLMNNNOTQLFLEHSLTANTTGTIT-- 297  
 QY 207 AGICGCG-LALALIFGALIFKWSHSEKIQNLISLANLPSSGLANAVAGIRSEENITY 265  
 DB 238 AGVCISVIVLALLLVITIAKKYFKKE-VQQLS-VSFSLQIKALQNAVEKVEQAEEDNY 355  
 QY 266 TIEENVYEE 275  
 DB 356 -IENSLYATD 364

## RESULT 10

Q8VBW0 PRELIMINARY; PRT; 305 AA.  
 AC Q8VBW0; 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE TIM2.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DA/2, and BALB/c; TISSUE=Spleen;  
 RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,  
 RA Freeman G.J., Umetsu D.T., Dekruyf R.H.;  
 RT "Tapr", a major T cell regulatory locus that controls the development  
 of airway hyperactivity, cosegregates with variants in a novel gene  
 family.";  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF399828; AAJ35773.1; -.  
 DR EMBL; AF399827; AAJ35772.1; -.  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 SQ SEQUENCE 305 AA; 33519 MW; 68E69370FE23C9D0 CRC64;

Query Match 17.2%; Score 273.5; DB 11; Length 305;  
 Best Local Similarity 28.0%; Pred. No. 8.1e-18;  
 Matches 84; Conservative 41; Mismatches 114; Indels 41; Gaps 9;

QY 12 LLLLLLTSSSEVYRAEVGNAYLPCFYTPAAGNLPVPCMGKACVFECCGNVLRD 71  
 DB 11 LILLPGAVESHNTAVQGLAGHPVLPCLYSTHNL-GGIVPMKMGSGEGRHSCTISLWTN 69  
 QY 72 ERDVNYW-TSRYWNGDFPKGDVSLTIENTVTLADSGIYCCRIQIPGIMNDEKFNK 130  
 DB 70 GYVTHORNSRYQKGNISEGNVSLTIENTVVGDDGVCVEVIRGAFHFDVYMLE-VK 127  
 QY 131 PAKTPAFTLORDTAAFPRLTTRG--HGAETQTLGSLPDINLQI 171  
 DB 128 PEISTSPPT--RPAATGPTTSTRSTHVPSTIVSTSTSPPTAHTETKREATTPTPDQ 185  
 QY 172 NLQI-----STLANELDRSLANDLRDSGA-----TIRIGIYAGICAGLALA 216  
 DB 186 TTAETVETLPSTPDMDNNTVSSDDPMDNNEVAPQKPKQNLKGFVVGSI--AALLIL 244  
 QY 217 LIFGALIFKWSHSEKIQNLISLANLPSSGLANAVAGIRSEENITYTIEENVYEE 276  
 DB 245 MLSTWVITRYVMKRSSESLSFVAFPISKIAGSPKVVETRCEDQVYIIEDTPYPREE 304

## RESULT 11

Q8RI83 PRELIMINARY; PRT; 305 AA.  
 AC Q8RI83; 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Hypothetical 33.5 Kda protein.  
 GN TIMD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025096; AAH25096.1; -.  
 DR MGD; MGI:2159681; Timd2.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.



DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Hepatitis A virus receptor.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97368416; PubMed=9225030;  
 RX Ashida M., Hamada C.;  
 RT "Molecular cloning of the hepatitis A virus receptor from a simian  
 cell line."  
 RL J. Gen. Virol. 78:1565-1569(1997).  
 DR EMBL, D85855; BAA21556.1; -.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR002965; P-rich extenan.  
 DR PRINTS: PR01217; PRICHEXTENS.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Receptor.  
 SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;

Query Match 15.7%; Score 248.5; DB 6; Length 460;  
 Best Local Similarity 22.5%; Pred. No. 3.6e-15;  
 Matches 103; Conservative 37; Mismatches 123; Indels 195; Gaps 13;

QY 4 HLPFCVLLLLLLLTSSSEVEYRAVGNAYLPCEYTPAPAGNLVPCWKGACPFVEEC 63  
 DB 2 HLOVVISLILHLADPVAADVAVGAGLSTLPKRN---GALISCMWNKGTCSVFSC 57  
 QY 64 GNVVLRTDERDVNY-WTSRYWLGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMNDEK 122  
 DB 58 PDGIWVTGHTVYTKRETRYKLGNLSDRDVSLTANTAVSDSGIYCCRVHSGWFMNKK 117  
 QY 123 FNLKVIIPAKYT----- 135  
 DB 118 ITISLKIIPRVTVIVRTVSTVPTTTTLPPTTTLPTTTLPTTTLPTTTLPTTTLPM 177  
 QY 136 -----PAFTLQDRF--TAAPRMLTTRGHGPAETQL----- 165  
 DB 178 TTTLPPTTTTPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 236  
 QY 166 -----GSLPDINLTQISTLANEL----- 183  
 DB 237 MTTLPPTTTTPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 296  
 QY 184 -----RDSRLANDLRDSGATI----- 199  
 DB 297 PPTPLPMONHEPVATSPSSPPAPETHPVTLGATRTQPTSPSYGTYDGDSTVTESSDG 356  
 QY 200 -----RIGIYIGAGICAG-LALALIFGALIFKWSHKEKIONLS 238  
 DB 357 LMNNQOTQLSPRHSQPMQNTTEGTY--AGVCISVLLAVLAVGVIAKKYFFKKE-IQQLS 413  
 QY 239 LLSLANLPPSGLANAVAGIRSEENIYTIENVEVEE 276  
 DB 414 -VSFSNHQKTLQNAVKKEVHAEDNIY-IENNLVYAMNQ 449

RESULT 15

ID Q96H15 PRELIMINARY; PRT; 378 AA.  
 AC Q96H15;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BC008988; AAH08988.1; -.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Hypothetical protein.  
 SQ SEQUENCE 378 AA; 41550 MW; 758B9E90CB28FEED CRC64;

Query Match 15.5%; Score 245.5; DB 4; Length 378;  
 Best Local Similarity 36.2%; Pred. No. 5.4e-15;  
 Matches 63; Conservative 25; Mismatches 77; Indels 9; Gaps 4;

QY 15 LLLLTSSSEVEYRAVGNAYLPCEYTPAPAGNLVPCWKGACPFVECGNVVLTQD-ER 73  
 DB 17 LYLFPVTSETVTEVLGHRVTLPCLYS-SWGHNSNSMCMGKQDCPYSGCKEALRTDGM 75  
 QY 74 DVNTWTSRYWLGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMNDEKFNKVIKPAK 133  
 DB 76 VTSRSKSAKYRLQGTIPRGDVSLLTINPESDSGYCCRIEVPGWENDKVKNVRLMQRAS 135  
 QY 134 VT---PAFTLQDRF--TAAPRMLTTRGHGPAETQLGSLPDINLTQISTLA 180  
 DB 136 TTTTATTTTTRRTTTTSTPTTTRQNTTTPALPTTTLPTTTLPTTTLPTTTLPT 189

Search completed: November 22, 2003, 05:43:24  
 Job time : 61.4416 secs

**THIS PAGE BLANK (USPTO)**



	Matches	301:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
Qy		1	MFSHLPFCVVLILLILLITRSSEVYRAEVGONATLPCTPTTAAAGNLVPVWGKACPV						60	
Db		1	MFSHLPFCVVLILLILLITRSSEVYRAEVGONATLPCTPTTAAAGNLVPVWGKACPV						60	
Qy		61	FECSNVVLRTERDQVNWYTSRYWMLNGDRKGDVSLTIENVTLADSGIYCCRIQIPGINMD						120	
Db		61	FECSNVVLRTERDQVNWYTSRYWMLNGDRKGDVSLTIENVTLADSGIYCCRIQIPGINMD						120	
Qy		121	EKNFLKLVIKPAKYTPATPLORDFPAAPRPMLTTRGHGAETQTSJPDINLTQISTIA						180	
Db		121	EKNFLKLVIKPAKYTPATPLORDFPAAPRPMLTTRGHGAETQTSJPDINLTQISTIA						180	
Qy		181	NELPDSRLANDLRDSGATIRIGIYAGACIAGALALAFGALIIPKWSHSEKIONLSLI						240	
Db		181	NELPDSRLANDLRDSGATIRIGIYAGACIAGALALAFGALIIPKWSHSEKIONLSLI						240	
Qy		241	SLANLPBEGLANAAVEGIRSEENIYTIENNYVEVEPEPNEYCYVSSROQSPQLGRFAM						300	
Db		241	SLANLPBEGLANAAVEGIRSEENIYTIENNYVEVEPEPNEYCYVSSROQSPQLGRFAM						300	
Qy		301	P 301							
Db		301	P 301							

## RESULT 2

ID Q8TDQ0 PRELIMINARY; PRT; 301 AA.

DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hepatitis A virus cellular receptor 2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP  
RP SEQUENCE FROM N.A.  
RA Zhang W., Man T., Li N., Gao X.;  
RT "Novel human hepatitis A virus cellular receptor.";  
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF251707; AA019100.1; -.  
DR Genew; HGNC:18437; HAVCAR2.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-1like.  
DR SMART; SM00409; IG\_1.  
DR PROSITE; PSS0835; IG\_LIKE; 1.  
DR  
KW Receptor.  
KW  
KW SEQUENCE 301 AA; 3363 MW; 7DDEA13F01D35878 CXC6;

Query Match	99.6%	Score 1581;	DB 4;	Length 301;
-------------	-------	-------------	-------	-------------

Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY	1	MFESHLPFCVLLMLLLLLLRSRSEVEBRAVQGNAYLPCYTPPAPNLVPVCKGKACVY	60
Db	1	MPSHLPFDCVLLLLLLLLLRSRSEVEBRAVQGNAYLPCYTPPAPNLVPVCKGKACVY	60
QY	61	FECGNVLTDERDVNYWTSRYWLNDFPKGDVSLTENVTLADSGIYCCRIQIPGIMND	120
Db	61	FECGNVLTDERDVNYWTSRYWLNDFPKGDVSLTENVTLADSGIYCCRIQIPGIMND	120
QY	121	EKENLKVYKPAKWTAPLQRPDFTPAFFRMLLTRHGGAETOTGLSLPDINTOISTLA	180
Db	121	EKENLKVYKPAKWTAPLQRPDFTPAFFRMLLTRHGGAETOTGLSLPDINTOISTLA	180
QY	181	NELRLANDLSDSGATIRIGIYGAGICAGLALLIFGALFFKYSHSKEKIQNLISL	240
Db	181	NELRLANDLSDSGATIRIGIYGAGICAGLALLIFGALFFKYSHSKEKIQNLISL	240

QY 241 SLNLTPEPSGLANVAAGIRSEENITYTIEENVYAEVEPEMYCYVSSRQSPQPLGCRFAM 300  
Db 241 SLNLTPEPSGLANVAAGIRSEENITYTIEENVYAEVEPEMYCYVSSRQSPQPLGCRFAM 300  
QY 301 P 301  
Db 301 P 301

### RESULT 3

Q8VIM0	PRELIMINARY;	PRT;	281 AA.
ID			
Q8VIM0			

DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)  
DE TIMD3.  
GN Mus musculus (Mouse).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
OX NCBI\_Taxid=10090;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2; TISSUE=Spleen;  
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Barsh G.S.,  
RA Freeman G.D., Umetsu D.T., Detrynff R.H.,  
RT "Apur, a major T cell regulatory locus that controls the developm  
RT of airway hyperreactivity, cosegregates with variants in a novel s  
RL family." (JUL-2001) to the EMBL/GenBank/DBD databases.  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBD databases.  
DR EMBL; AF399831; AL33776.1; -.  
DR MGI; MGI:2159682; T1md3.  
DR InterPro; IPR003589; IG.  
DR InterPro; IPR007110; IG-1like.  
DR InterPro; IPR000719; Prot\_kinase.  
DR SMART; SMO0409; IG: 1.  
DR PROSITE; PSS0835; IG\_LIKE; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
SQ SEQUENCE 281 AA; 30934 MW; C0349E4BD0E5761D CRC64;

Query Match	Score	DB	Length
55.8%	885	11	281

Matches 187; Conservative 26; Mismatches 67; Indels 12; Gaps 4;

QY	1	MESHLPFCVLLLLLLLLTRSSSEVYRAVEQONALPFCYTPAAGNLVPCWKGACPV	60
QY	1	MESHLPFCVLLLLLLLLTRSSSEVYRAVEQONALPFCYTPAAGNLVPCWKGACPV	60
Db	1	MFSGITLNCVLLLLLOLLARSLLENAYFEVGNAYLPSTYLSLTPGALVPMCMGFGCPW	60
QY	61	FECCGVVLRTERDENVYVW-TSRWYLMNGDFPRKGDVSLTTEENVTLADSGIYCCRIOIQTGMN	119
Db	61	SOCTNELIRTERNVITYOKSSRYOLKGLNKGDSLLIKNVTLDHGYCCRIQFPGMLN	120
QY	120	DEKENLKLIVIPAKVTPAPTLQORDFTAAFPMLITRGHPAETQTLGSLPDINLTQISTL	179
Db	121	DKKLEIKLDIAAKVTPAQTAHGSSTASPRTLTTERNG-SEPTQTLVTLHNNNGRIKISTW	179
QY	180	ANELRDSRLANDLRDSGATIRIGIYGAGICQGLALALIFGALIFKMTYSHSKKIQONISL	239
Db	180	ADE-----IKSGETIRTAIHIGVGSAGLTLALIGVLILKWSCKKKKJSSLSL	230
QY	240	ISLANLPFSGLANNAVEGIRSEENITYTEENVYFVEEENEYCYSSQOPS	291
Db	231	ITLANLPFGLANAGAVIRISEENITYTEENVYEVENSENEYCYVNS-QOPS	291
RESULT 4			
ID	Q8WM60	PRELIMINARY;	PRT; 142 AA.
DT	01-MAR-2002	(TREMBlrel). 20.	Created)
DT	01-MAR-2002	(TREMBlrel). 20.	Last annotation update)
DT	01-MAR-2003	(TREMBlrel). 23.	Last annotation update)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:03:41 / Search time 58.4416 Seconds  
(without alignments)  
1329.085 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MESHLPFDCCVLLLLLLLLTR.....CYVSSRQPSQPIGRFAMP 301

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPREMBL\_23.\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	301	Q96K94	Q96K94 homo sapien
2	1581	99.6	301	Q8TDC0	Q8TDC0 homo sapien
3	885	55.8	281	Q8VIM0	Q8VIM0 mus musculu
4	718	45.2	142	Q8WV60	Q8WV60 homo sapien
5	336	21.2	307	054947	054947 rattus norv
6	333	21.0	282	11 Q8VIM1	Q8VIM1 mus musculu
7	323.5	20.4	305	11 Q8VIM2	Q8VIM2 mus musculu
8	322	20.3	359	4 Q8VIM2	Q8VIM2 mus musculu
9	316.5	19.9	364	4 Q8VIM2	Q8VIM2 mus musculu
10	273.5	17.2	305	11 Q8VIM0	Q8VIM0 mus musculu
11	263.5	16.6	305	11 Q8VIM0	Q8VIM0 mus musculu
12	256	16.1	451	6 Q95144	Q95144 cercopithec
13	248.5	15.7	362	11 Q8CIC7	Q8CIC7 mus musculu
14	248.5	15.7	460	6 Q18984	Q18984 cercopithec
15	245.5	15.3	378	4 Q96H15	Q96H15 homo sapien
16	243	15.3	474	6 Q46597	Q46597 cercopithec

17	241	15.2	478	6 Q46598	Q46598 cercopithec
18	137.5	8.7	258	4 Q9UMT2	Q9UMT2 homo sapien
19	135	8.5	276	4 Q95944	Q95944 homo sapien
20	133.5	8.4	258	4 Q9H563	Q9H563 homo sapien
21	132.5	8.3	270	4 Q9UMT1	Q9UMT1 homo sapien
22	131.5	8.3	336	4 Q8WV55	Q8WV55 homo sapien
23	131.5	8.3	352	11 Q9UMT6	Q9UMT6 mus musculu
24	131.5	8.3	365	11 Q9DBJ8	Q9DBJ8 mus musculu
25	131.5	8.3	523	4 Q00480	Q00480 homo sapien
26	131	8.3	276	4 Q9H562	Q9H562 homo sapien
27	130	8.2	319	6 Q9TUB0	Q9TUB0 canis fam11
28	130	8.2	344	11 Q9R067	Q9R067 rattus norv
29	130	8.2	358	11 Q9R066	Q9R066 rattus norv
30	128.5	8.1	270	4 Q9H564	Q9H564 homo sapien
31	125.5	7.9	319	6 Q9TUB9	Q9TUB9 sus scrofa
32	124.5	7.8	344	4 Q9UKV4	Q9UKV4 homo sapien
33	123	7.8	200	4 Q8WV77	Q8WV77 homo sapien
34	121	7.6	334	4 Q96AV7	Q96AV7 homo sapien
35	121	7.6	527	4 Q00475	Q00475 homo sapien
36	120	7.6	365	6 Q8WV23	Q8WV23 bos taurus
37	119.5	7.5	455	4 Q9UIR0	Q9UIR0 homo sapien
38	118	7.4	252	4 Q8WV16	Q8WV16 homo sapien
39	116.5	7.3	287	11 Q9D7B8	Q9D7B8 mus musculu
40	116	7.3	259	4 Q95532	Q95532 homo sapien
41	115	7.2	318	13 Q91664	Q91664 xenopus lae
42	114	7.2	789	11 Q8CJ59	Q8CJ59 mus musculu
43	112.5	7.1	430	4 Q8N4F1	Q8N4F1 homo sapien
44	111.5	7.0	319	11 Q922D5	Q922D5 mus musculu
45	111.5	7.0	529	4 P78408	P78408 homo sapien

## ALIGNMENTS

RESULT 1

Q96K94 PRELIMINARY; PRT; 301 AA.

AC Q96K94; 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE Hypothetical protein F141428 (Putative kidney injury molecule-3).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OK NCBI\_taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE=Embryo;

RA Isovai T., Ota T., Hayashi K., Sugiyama T., Ofuku T., Suzuki Y., Nishikawa T., Negai K., Sugano S., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.; "NEBO human cDNA sequencing project."

RT Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA TISSUE=Hepatoma;

RA Kuen E.W., Ichimura T., Bonventre J.V.; "A homolog to human kidney injury molecule-1 is expressed in hepatoma cells."

RT Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.

RU EMBL; AK077334; BAB5044.1; -

DR EMBL; AY069944; BAB5040.1; -

DR InterPro; IPR007110; IG-1like.

DR PROSITE; PS50835; IG\_LIKE; 1.

KV Hypothetical protein\_

SQ SEQUENCE 301 AA; 33351 MW; 487787D9E40A4161 CRC64;

Query Match 100.0%; Score 1587; DB 4; Length 301;

Best local Similarity 100.0%; Pred. No. 3.7e-144;



**THIS PAGE BLANK (USPTO)**



```

FT FT isoform 5 and isoform 7).
FT FT /FTId=VSP 002920.
FT FT G-> EQWVTVYNSDVHNNSTASDNNRLG (1n isoform
FT FT 8 and isoform 9).
FT FT /FTId=VSP 002921.
FT FT G-> VHGKGVGVLDQIWLSDQDNEEQVMTVNSDVH
FT FT NNSTASDNNRLG (1n isoform 10 and isoform
FT FT 12).
FT FT /FTId=VSP 002922.
FT FT Missing (1n isoform 4).
FT FT /FTId=VSP 002923.
SQ SEQUENCE 818 AA, 91736 MW, D1BA39E2092B2152 CRC64;

Query Match 6.2%; Score 99; DB 1; Length 818;
Best Local Similarity 19.9%; Pred. No. 0.65;
Matches 59; Conservative 41; Mismatches 89; Indels 108; Gaps 12;

QY 87 DPKRQDVSLITENVTLADSGIY--CCRIQIPG-INNDEKF-----NKLVIKP----- 131
DB 241 DTSKNPASLTITKNVSMDSGLWISCAVENIVGEVOTSALIVFAPANITFIESPDPDHW 300
QY 132 ----AKYTPAPTLORDFTAAPR-----M 151
DB 301 CIPFVTKGNPKFTLQMFYEGALINSEYICTKHIVINOSHYGCLQDNPHTLNNGAYTL 360
QY 152 LTRGHGPAETQT-----LGSLPDINLTQISTLANELRDSR-----LAN 190
DB 361 LAKNEYGEDEKRVDAHFMFSVPGDGGPIYDPVYEYETTPNDLGGTNNNSQITSPDVSN 420
QY 191 DLDRSGATIRIGIYAGAGICAGIALALIFGALIFKMYSHSKKIKONLSI----- 240
DB 421 KENEDSIYVVVVGIAALVCTGLVIMLI-----ILKFGRSKFGKMGKPPSSVISNDDSDASP 476
QY 241 ----SLANLPPS--GLANAVAGIRSEENITYIENYVEVEPEHYCYVSSROOP 290
DB 477 LHHISNGSNTSPSSSGGPDVAITGM-----TKIFVIENP-QYFITTSQLKP 522

RESULT 13
MOG_RAT STANDARD; PRT; 245 AA.
ID 063145;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Mathieu J.-M.;
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily.";
RL J. Neurosci. Res. 33:177-187 (1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.;
RA Pontarotti P., Roessel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
RN [3]
RP STRUCTURE BY NMR OF 62-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouze-Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.;

```

```

RT RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55.";
RT Eur. J. Biochem. 246:59-70 (1997).
CC CC -1- FUNCTION: MINOR COMPONENT OF THE MYELIN SHEATH. MAY BE INVOLVED IN
CC CC COMPLETION AND/OR MAINTENANCE OF THE MYELIN SHEATH AND IN CELL-
CC CC CELL COMMUNICATION.
CC CC -1- SUBUNIT: MAY FORM HOMODIMERS.
CC CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE CNS, WHERE IT IS
CC CC LOCALIZED ON THE SURFACE OF MYELIN AND OLIGODENDROCYTE CYTOPLASMIC
CC CC MEMBRANES.
CC CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC CC ACTIVE MYELINATION.
CC CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG
CC CC SUBFAMILY.
CC CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC CC -1- CAUTION: DO NOT CONFUSE MYELIN-OLIGODENDROCYTE GLYCOPROTEIN (MOG)
CC CC WITH OLIGODENDROCYTE-MYELIN GLYCOPROTEIN (OMG).
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; M99485; AAA41628.1; -
DR DR EMBL; L21995; AAF74786.1; -
DR DR PIR; B47712; B47712.
DR DR InterPro; IPR007110; IG-1-like.
DR DR InterPro; IPR003006; IG_MHC.
DR DR InterPro; IPR003596; IG_V.
DR DR Pfam; PF00047; Ig_1.
DR DR SMART; SM00406; IGV_1.
DR DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT FT SIGNAL 1 27
FT FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
FT FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 156 176 POTENTIAL.
FT FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 209 229 POTENTIAL.
FT FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
FT FT DOMAIN 30 139 IG-LIKE.
FT FT DISULFID 51 125 POTENTIAL.
FT FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 245 AA; 27881 MW; C97F8AD60D6A32B4 CRC64;

Query Match 6.2%; Score 98.5; DB 1; Length 245;
Best Local Similarity 27.2%; Pred. No. 0.16;
Matches 37; Conservative 18; Mismatches 66; Indels 15; Gaps 3;

QY 5 LPFDCVILLILLILTRSEVEY-----RAEYQNAVLPCEFTPAAPGNTLVPCWCK 55
DB 9 LPSCLSLILLILQSRSAVAGFVYIGPHPIRALVGDABALPCRISPKNATGMVEGMYR 68
QY 56 GACPVEEGNAVLTDERDV---NYTSTRYWLNGDFPKGVSLTIENTVTLADSGIYCCR 111
DB 69 S--PFSRVVHLRYRNGKDDAQAQAEYRGRTLELKEISIGEGVALRIQNVRFSDGEGYTCF 126
QY 112 IQIPGIMNDEKFNKL 127
DB 127 FRDHSYGEAAVELKV 142

RESULT 14
BUTY_HUMAN STANDARD; PRT; 526 AA.
ID BUTY_HUMAN
AC Q13410;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

CC -1- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),  
CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR  
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE  
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN  
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-  
CC GAMMA-1 (BY SIMILARITY).  
CC  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC  
CC -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).  
CC  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=12;  
CC Comment=Additional isoforms seem to exist;.  
CC Name=1; Synonyms=Alpha-FL;  
CC IsoId=Q91987-1; Sequence=Displayed;  
CC Name=2; Synonyms=Beta-FL;  
CC IsoId=Q91987-2; Sequence=VSP\_002914;  
CC Name=3; Synonyms=ED;  
CC IsoId=Q91987-3; Sequence=VSP\_002915;  
CC Name=4; Synonyms=JD;  
CC IsoId=Q91987-4; Sequence=VSP\_002923;  
CC Name=5; Synonyms=J1;  
CC IsoId=Q91987-5; Sequence=VSP\_002920;  
CC Name=6; Synonyms=Alpha-T1;  
CC IsoId=Q91987-6; Sequence=VSP\_002918, VSP\_002919;  
CC Name=7; Synonyms=J1+T1;  
CC IsoId=Q91987-7; Sequence=VSP\_002918, VSP\_002919, VSP\_002920;  
CC Name=8; Synonyms=J2+T1;  
CC IsoId=Q91987-8; Sequence=VSP\_002918, VSP\_002919, VSP\_002921;  
CC Name=9; Synonyms=ED J2+T1;  
CC IsoId=Q91987-9; Sequence=VSP\_002915, VSP\_002918, VSP\_002919,  
CC VSP\_002921;  
CC  
CC Name=10; Synonyms=J1+J2+T1;  
CC IsoId=Q91987-10; Sequence=VSP\_002918, VSP\_002919, VSP\_002922;  
CC Name=11; Synonyms=T3;  
CC IsoId=Q91987-11; Sequence=VSP\_002916, VSP\_002917;  
CC Name=12; Synonyms=ED J1+J2+T1;  
CC IsoId=Q91987-12; Sequence=VSP\_002915, VSP\_002918, VSP\_002919,  
CC VSP\_002922;  
CC  
CC -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.  
CC  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN  
CC RECEPTOR SUBFAMILY.  
CC  
CC -1- SIMILARITY: Contains 2 leucine-rich (LRR) repeats.  
CC  
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
CC  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER LBU-144 OR MET-188 IS THE  
CC INITIATOR OF ISOFORM 2.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/isb-sib.ch>).  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; X77251; CA54468.1; -  
CC EMBL; X77252; CA54469.1; -  
CC EMBL; X74109; CA52210.1; -  
CC PIR; S59939; S44098.  
CC HSSP; P06213; IIRK.  
CC InterPro; IPR007110; IG-like.  
CC InterPro; IPR003598; IG\_C2.  
CC InterPro; IPR003006; IG\_MHC.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR000483; LRR\_Cterm.  
CC InterPro; IPR000372; LRR\_Kterm.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR002011; RTKinaseII.  
CC InterPro; IPR001245; Ty\_pkinase.  
CC Pfam; PF000047; Ig\_1.  
CC Pfam; PF00560; LRR; 1.  
CC Pfam; PF01463; LRRCT; 1.

[illegible]

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malik J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A., Smalls D.E.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Probable immunoglobulin-like cell surface receptor.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=Q9P1W8-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9P1W8-2; Sequence=VSP\_007027;  
 CC Note=No experimental confirmation available;  
 CC Name=3;  
 CC IsoId=Q9P1W8-3; Sequence=VSP\_007028;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Detected in liver, and at very low levels in  
 CC brain, heart, lung, pancreas, kidney, placenta and skeletal  
 CC muscle.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL AB042624; BAA95692.1; -;  
 CC EMBL AL138804; CAC00474.1; -;  
 CC EMBL BC020629; AAH20629.1; ALT\_INIT.  
 CC Genew; HGNC:15757; SIRPB2.  
 CC MIM: 605466; -;  
 CC DR GO: GO:0007267; P: cell-cell signaling; TAS.  
 CC DR GO: GO:0007242; P: intracellular signaling cascade; TAS.  
 CC DR GO: GO:0008285; P: negative regulation of cell proliferation; TAS.  
 CC DR InterPro: IPR003597; IG\_C1.  
 CC DR InterPro: IPR003006; IG\_MHC.  
 CC DR Pfam: PF00047; Igr\_3.  
 CC DR SMART; SM00407; IGc1\_2.  
 CC DR SMART; SM00406; IGV\_1.  
 CC DR PROSITE; PS50835; IG\_LIKE; 3.  
 CC DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 CC Repeat: Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;  
 KW Alternative splicing;  
 FT SIGNAL 1 28  
 FT CHAIN 29 387  
 FT DOMAIN 29 360  
 FT TRANSMEM 361 383  
 FT DOMAIN 384 387  
 FT DOMAIN 29 137  
 FT DOMAIN 146 245  
 FT DOMAIN 252 340  
 FT DISULFID 53 119  
 FT DISULFID 168 226  
 FT DISULFID 221 329  
 FT CARBOHYD 243 243  
 FT CARBOHYD 268 268  
 FT CARBOHYD 309 309  
 FT CARBOHYD 317 317

FT VARSPLIC 1 33 Missing (in isoform 2).  
 FT FTID=VSP\_007027.  
 FT VARSPLIC 144 360 Missing (in isoform 3).  
 FT FTID=VSP\_007028.  
 FT CONFLICT 263 263 A -> V (IN REF. 2).  
 FT CONFLICT 286 286 L -> S (IN REF. 2).  
 FT SEQUENCE 387 AA; 42495 MW; PFE20C9F6E064B CRC64;  
 SQ  
 Query Match 6.3%; Score 100.5; DB 1; Length 387;  
 Best Local Similarity 25.5%; Pred. No. 0.19; Indels 59; Gaps 15;  
 Matches 71; Conservative 37; Mismatches 111;  
 QY 4 HLPDPCVLLLLLLLTTR-SSEVER- - - - -AEVQONAYLPFCYPPAAGNLPVPCW 53  
 DB 9 HPGPFLITLLTGTEVAGEELQMIPEKLLVTGKTAHTHCTVSLP--VGPVLW 66  
 QY 54 GKACPVCEGCVVLRTERDVNTSYWMLNGPRK--DVSITENVTLADGIYCC 110  
 DB 67 FRGVGPGR- - - - -LIVNKGGRF--PRVTVSIDLTKNNMDFSIISITPADVGTYYC 119  
 QY 111 RIQIPGINNDEK- - - - -NLKVIKPAK- - - - -VTPAPTLQRFDTAFPRMLTTRGHGPAE 161  
 DB 120 VKFRKSPENVEFSGPETERALAKPAPVVLGPAATTPBHTVSF- - - - -TCSEHG- - - - - 172  
 QY 162 TQTLGSLPDILNLTQISTLANELRDSRGLANDRDSGATIRIGYIGAGICAGIALALIRGA 221  
 DB 173 - - - - -FSPRDITLTKMKK-CNELSD- - - - -POTNVDPICQSAVYSIRSTARV- - - - - 213  
 QY 222 LIFKWSHSHKEKIONLSISLANLPPSGLANAVAGIR 259  
 DB 214 VLDPWDVRS-OVICEVAHVTLQGDPLRGTA-N-LSEAIR 249  
 RESULT 12  
 TRKB\_CHICK STANDARD; PRT; 818 AA.  
 AC Q91987; Q91010;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE BDNF/NF-3 growth factors receptor precursor (EC 2.7.7.112) (TrkB  
 DE tyrosine kinase) (TrkB-B).  
 GN GN  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 10).  
 RC TISSUE=Brain;  
 RX MEDLINE=95047511; PubMed=7959025;  
 RA Vith N., Erdmann K., Heumann R.;  
 RT "Cloning and sequence analysis of a cDNA encoding a novel truncated  
 RT form of the chicken TrkB receptor.";  
 RL Gene 149:383-384(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=94116452; PubMed=8287802;  
 RA Dechant G., Biffo S., Okazawa H., Kolbeck R., Pottgiesser J.,  
 RA Barde Y.A.;  
 RT "Expression and binding characteristics of the BDNF receptor chick  
 RT trkB.";  
 RL Development 119:545-558(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE FROM N.A.  
 RP (ISOFORMS 2; 3; 4; 5; 6; 7; 8; 9; 10; 11 AND 12).  
 RX MEDLINE=96370546; PubMed=8774442;  
 RA Garner A.S., Mengay H.J., Boeshore K.L., Xie X.Y., Voci J.M.,  
 RA Johnson J.E., Lange T.H.;  
 RT "Expression of TrkB receptor isoforms in the developing avian visual  
 RT system.";  
 RL J. Neurosci. 16:1740-1752(1996).







FT DOMAIN 446 540 FIBROECTIN TYPE-III 1.  
 FT DOMAIN 543 639 FIBROECTIN TYPE-III 2.  
 FT DOMAIN 643 744 FIBROECTIN TYPE-III 3.  
 FT DOMAIN 839 1118 PROTEIN KINASE.  
 FT NP BIND 845 853 ATP (BY SIMILARITY).  
 FT BINDING 870 870 ATP (BY SIMILARITY).  
 FT ACT SITE 979 979 BY SIMILARITY.  
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MOD RES 1007 1007 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 1138 AA; 125089 MW; 3842B3678C58A1 CRC64;  
 Query Match 6.8%; Score 108; DB 1; Length 1138;  
 Best Local Similarity 23.6%; Pred. No. 0.15;  
 Matches 47; Conservative 32; Mismatches 46; Gaps 8;  
 QY 17 LITRSEVERAEVGNAYLPCEYTPAPGNTLVPCWGKACPFECGNAVLTDERDVN 76  
 DB 353 IINMASELEFNFLETMPR--INC---AAAGNPPV---RGSIELRKPDDGLSTVAIV- 402  
 QY 77 WYTSRWMLNGDRKGVSLIENVTLADSGIYCCRIQIGIMNDEKENLKVIRKAKTP 136  
 DB 403 -----EPEKTAEEFEPRLVADSGFWECKRSTSGGQDSRRFKVNV---KVP 447  
 QY 137 APTLORDFTAEPRLTTR-----GHGPAETQTGLSLPDIMLTOSTLANEL 183  
 DB 448 VP-----LAAPRLITKSRQVLVSPVSGDGIISTVRLHYRQDSTMWSTIVDP 500  
 QY 184 RDSRLANDLR-DSGATIRI 201  
 DB 501 SENVTLMNLRPKGYSVRV 519  
 RESULT 8  
 HEMA\_RACVI STANDARD; PRT; 310 AA.  
 ID HEMA\_RACVI 000716;  
 AC 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Hemagglutinin precursor.  
 GN HA.  
 OS Raccoon poxvirus (RCN).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxID=10256;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=92410621; PubMed=1529542;  
 RA Cavallaro K.F., Esposito J.J.;  
 RT "Sequences of the raccoon poxvirus hemagglutinin protein."  
 RL virology 190:434-439(1992).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: MAY BE PHOSPHORYLATED AND SULFATED. O-LINKED GLYCOSYLATION  
 CC MAY BE REQUIRED FOR HEMAGGLUTINATION AND HEMADSORPTION ACTIVITIES  
 CC OF INFECTED CELL MEMBRANS.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; M94169; AAA47231.1; -.  
 CC PIR; A43381; HNVZRA.  
 CC InterPro; IPR007110; IG-1like.  
 CC InterPro; IPR003599; IG.

DR InterPro: IPR003006; IG\_MHC.  
 DR Pfam: PF00047; IG\_1.  
 DR SMART; SM00409; IG\_1.  
 DR PROSITE; PS50835; IG\_LINE; 1.  
 KW Hemagglutinin; Envelope protein; Late protein; Glycoprotein; Signal;  
 KW Transmembrane; Immunoglobulin domain.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 310 HEMAGGLUTININ.  
 FT DOMAIN 17 275 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 276 300 POTENTIAL.  
 FT DOMAIN 301 300 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 19 121 IG-LIKE V-TYPE.  
 FT DISULFID 36 105 POTENTIAL.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 310 AA; 34189 MW; 7C25F5C19C7B11B9 CRC64;  
 Query Match 6.7%; Score 107; DB 1; Length 310;  
 Best Local Similarity 23.2%; Pred. No. 0.038;  
 Matches 53; Conservative 36; Mismatches 81; Indels 58; Gaps 10;  
 QY 10 VILLILLILLITRSEVERAEVGNAYLPCEYTPAPAGNLVPCWGKACPFECGNAVLR 69  
 DB 8 ILLIIVVTKKHPQISKQGDADRLSC-NRNNTHGLVMSWK-----KPSIILL 60  
 QY 70 TDERDVNWTSRWMLNGDRKGVSLIENVTLADSGIYCCRIQIGIM 118  
 DB 61 AAKNDVVVF-----DDVTADKVSVDSPDTLITIKSLTSDAGYLCAFITTS-T 112  
 QY 119 ND-----EKENLKVIRKAKT-----PAPTLORDFTAEPRLTTRGHGPAET-QTL 165  
 DB 113 NDTKIDVEEYFIDLVNPNANVTIDALISGTTODIISH-----TERQHSDDTICTS 167  
 QY 166 GSLPDIMLTOSTLANELRDS-----RLANDLDSGATI 199  
 DB 168 ESTTQISTSESTSSQISTSESTSSGYVEDDTQVNVTTDTTDSPTI 215  
 RESULT 9  
 C166\_CARAU STANDARD; PRT; 555 AA.  
 ID C166\_CARAU 090304;  
 AC 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE CD166 antigen homolog precursor (Neurolin) (DM-GRASP homolog).  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Retina;  
 RC laeasing U.;  
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 30-552 FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina;  
 RX MEDLINE=94299040; PubMed=8026643;  
 RA laeasing U., Giordano S., Stecher B., Lottspeich F., Stuermer C.A.O.;  
 RT "Molecular characterization of fish neurolin: a growth-associated  
 RT cell surface protein and member of the immunoglobulin superfamily in  
 RT the fish retinotectal system with similarities to chick protein  
 RT DM-GRASP/SC-1/BEN."  
 RL DM-GRASP/SC-1/BEN."  
 CC -1- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY  
 CC NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: PRESENT ON ALB RETINAL GANGLION CELLS (RGCS)  
 CC AND THEIR AXONS (IN EMBRYO). ABSENT FROM MATURE AXONS ALONG MOST



RA Moritz R.L., Ji H., Heath J.K., White S.J., Welt S., Old L.J.,  
 RA Simpson R.J.;  
 RT "Characterization of posttranslational modifications of human A33  
 RT antigen, a novel palmitoylated surface glycoprotein of human  
 RT gastrointestinal epithelium.";  
 RT Biochem. Biophys. Res. Commun. 236:682-686 (1997).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL  
 CC EPITHELIUM AND IN 95% OF COLON CANCERS.  
 CC -1- PTM: N-GLYCOSYLATED. CONTAINS APPROXIMATELY 8 KDA OF N-LINKED  
 CC CARBOHYDRATE.  
 CC -1- PTM: PALMITOYLATED.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain;  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U79725; AAC0957.1; -  
 CC Genew; HGNC:4445; GPA33.  
 DR DR MIM; 602171; -  
 DR GO; GO:0005888; C:proteoglycan integral to plasma membrane; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS00335; IG\_LIKE; 2.  
 DR Immunoglobulin domain; lipoprotein; Palmitate; Glycoprotein;  
 KM Transmembrane; Signal; Antigen.  
 FT SIGNAL 1 21  
 FT CHAIN 22 319  
 FT DOMAIN 22 235  
 FT TRANSMEM 236 256  
 FT DOMAIN 257 319  
 FT DOMAIN 22 134  
 FT DOMAIN 140 227  
 FT DOMAIN 140 261  
 FT DISULFID 43 117  
 FT DISULFID 146 232  
 FT CARBOHYD 112 112  
 FT CARBOHYD 200 200  
 FT CARBOHYD 223 223  
 SQ SEQUENCE 319 AA; 35632 MW; 9BFC7AFAF45C2408E CRC64;  
 Query Match 7.1%; Score 112.5; DB 1; Length 319;  
 Best Local Similarity 22.7%; Pred. No. 0.013;  
 Matches 67; Conservative 47; Mismatches 108; Indels 73; Gaps 15;  
 QY 27 RAEVGNAYLPCEYTPAAGNLPVCGKACPFECGVVLTDERDV-----NY-- 77  
 DB 32 RASQSGSVLPCTYHSTSSREGLQMDK-----LLLHTERVVWIPRNNYH 81  
 QY 78 ---WTSRYLNGDFKGDVSLTENTVTLDSGYCCRIQPGIANDKFNLK-----LV 128  
 DB 82 GELYKRVASINNAEODASITTDLTMA-DNGTECVS--SI-MSDLEGNTSRVLLVLI 138  
 QY 129 IKPAKTPAFLORDTAAFPRLVTRGH-----GPARTQTL 165  
 DB 139 VPPSK--PEGGLEGTITGNIOITQSGEGSPTPOYSWKRYNIINOBLAOPASGQPV 196  
 QY 166 GSLPDIINTL---QISTLANELDRSLAN-DLDSGATTIRIGIYAGICAGIALAL-IF 219  
 DB 197 -SLKNISTDTSGYICTSNE-EGTQFCNTTAVRSPNMVALYV--GIAVGVALVLI 252

QY 220 GALLFKWTSKSEKIQNLISLISLANTL-----PPSLANAVAEGRSENIITYEE 269  
 DB 253 GIIVYCCCCRGKDNDTEDEKEDARPAREAYEPPPOLRELRSB--REEDDYROEE 305  
 RESULT 6  
 ID PVR2\_MOUSE STANDARD; PRT; 530 AA.  
 AC P32507; 062096;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Poliovirus receptor related protein 2 precursor (Murine herpesvirus  
 GN entry protein B) (mHvrb) (Nectin 2) (Poliovirus receptor homolog).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
 RX MEDLINE=92219365; PubMed=1560525;  
 RA Morrison M.B., Racanelli V.R.;  
 RT "Molecular cloning and expression of a murine homolog of the human  
 RT poliovirus receptor gene.";  
 RL J. Virol. 66:2807-2813 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM BETA).  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=94179228; PubMed=8132569;  
 RA Aoki J., Koike S., Ise I., Sato-Toshida Y., Nomoto A.;  
 RT "Amino acid residues on human poliovirus receptor involved in  
 RT interaction with poliovirus.";  
 RL J. Biol. Chem. 269:8431-8438 (1994).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99214397; PubMed=10196354;  
 RA Shukla D., Rowe C.L., Dong Y., Racanelli V.R., Spear P.G.;  
 RT "The murine homolog (Mpr) of human herpesvirus entry protein B (Hvrb)  
 RT mediates entry of pseudorabies virus but not herpes simplex virus  
 RT types 1 and 2.";  
 RL J. Virol. 73:4493-4497 (1999).  
 CC -1- FUNCTION: RECEPTOR FOR ALPHAHERPESVIRUS (SUCH AS MURINE HSV) ENTRY  
 CC INTO CELLS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Beta;  
 CC IsoId=P32507-1; Sequence=Displayed;  
 CC Name=Alpha;  
 CC IsoId=P32507-2; Sequence=VSP 002630; VSP 002631;  
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, SPLEEN, KIDNEY, HEART AND  
 CC LIVER.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M60206; AAA39734.1; -  
 CC EMBL; D26107; BA05103.1; -  
 DR PIR; A38211; HLMSP3.  
 DR PIR; A53437; A53437.  
 DR MGD; MGI:97822; Pvr12.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.

RA Ling V., Dunnesi-Joannopoulos K.;  
 RT "G150 molecules and uses thereof."  
 RL Patent number WO0121796, 29-MAR-2001.  
 CC -I- FUNCTION: LIGAND FOR THE T-CELL-SPECIFIC CELL SURFACE RECEPTOR  
 CC ICOS. ACTS AS A COSTIMULATORY SIGNAL FOR T-CELL PROLIFERATION AND  
 CC CYTOKINE SECRETION; INDUCES ALSO B-CELL PROLIFERATION AND  
 CC DIFFERENTIATION INTO PLASMA CELLS. COULD PLAY AN IMPORTANT ROLE IN  
 CC MEDIATING LOCAL TISSUE RESPONSES TO INFLAMMATORY CONDITIONS, AS  
 CC WELL AS IN MODULATING THE SECONDARY IMMUNE RESPONSE BY CO-  
 CC STIMULATING MEMORY T CELL FUNCTION. DURING PREGNANCY, MAY FUNCTION  
 CC TO SKEW THE CYTOKINE OF MATERNAL T-CELLS TOWARD IMMUNOPROTECTIVE  
 CC TH2 PHENOTYPE.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=Q9JH8-1; Sequence=Displayed;  
 CC Name=2; Synonyms=B;  
 CC IsoId=Q9JH8-2; Sequence=VSP\_002521;  
 CC -I- TISSUE SPECIFICITY: ISOFORM 1 HIGHEST EXPRESSION IN LYMPHOID  
 CC TISSUES, SUCH AS SPLEEN (MOSTLY IN THE MARGINAL ZONE), LYMPH NODES  
 CC (PARTICULARLY IN THE CORTX AND IN BOTH PRIMARY AND SECONDARY  
 CC FOLLICLES), THYMUS (PREDOMINANTLY IN THE MEDULLA) AND PETER'S  
 CC POLICES (MOSTLY IN THE FOLLICLES), LOWER LEVELS IN MANY  
 CC NON-LYMPHOID TISSUES, SUCH AS BRAIN, HEART, KIDNEY, LIVER, LUNG,  
 CC SKELETAL MUSCLE AND TESTIS. PRESENT ON FRESHLY ISOLATED SPLENIC B-  
 CC CELLS, T-CELLS, DENDRITIC CELLS AND MACROPHAGES. THE EXPRESSION OF  
 CC ISOFORM 2 IS RESTRICTED TO HEART, SPLEEN AND KIDNEY.  
 CC -I- DEVELOPMENTAL STAGE: DETECTED EARLY IN EMBRYOGENESIS. IN THE YOLK  
 CC SAC AT 11.5 AND 12.5 DPC AND, TO A LESSER EXTENT, IN THE LIVER AT  
 CC 14.5 DPC.  
 CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. B7N/MOG  
 CC SUPERFAMILY.  
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-stb.ch/announce/>  
 CC or send an email to [license@1sb-stb.ch](mailto:license@1sb-stb.ch)).  
 CC -----  
 CC EMBL: AF216747; AAF45149.1; -;  
 CC EMBL: AF198027; AAF34738.1; -;  
 CC EMBL: AX100591; CAC36463.1; -;  
 CC EMBL: AX100593; CAC36464.1; -;  
 CC EMBL: AF394451; AAK77544.1; -;  
 CC MGD: MGI:1354701; Icosl.  
 CC DR GO: GO:0005576; C:extracellular; IDA.  
 CC DR GO: GO:0016021; C:integral to membrane; NAS.  
 CC DR GO: GO:0003793; F:defense/immunity protein activity; NAS.  
 CC DR GO: GO:0005102; F:receptor binding activity; TAS.  
 CC DR GO: GO:0045190; P:isotype switching; NAS.  
 CC DR GO: GO:0042104; P:positive regulation of activated T-cell pro. . .; TAS.  
 CC DR GO: GO:0045404; P:positive regulation of interleukin-4 biosyn. . .; IMP.  
 CC DR GO: GO:0007165; P:signal transduction; NAS.  
 CC DR GO: GO:0042110; P:t-cell activation; NAS.  
 CC DR InterPro: IPR007110; IG-like.  
 CC DR InterPro: IPR003599; IG.  
 CC DR InterPro: IPR003066; IG\_MHC.  
 CC DR Pfam: PF00047; Ig\_1.  
 CC DR SMART: SM00409; Ig\_1.  
 CC DR PROSITE: PS50835; IG\_LIKE; 2.  
 CC KW B-cell activation; Immune response; Glycoprotein;  
 CC Immunoglobulin domain; Signal; Transmembrane; Multigene family;  
 CC Alternative splicing.  
 CC FT SIGNAL 1 46 BY SIMILARITY.  
 CC FT CHAIN 47 322 ICOS LIGAND.  
 CC FT DOMAIN 47 277 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSMEM 278 298 POTENTIAL.

FT DOMAIN 299 322 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 49 149 IG-LIKE V-TYPE.  
 FT DOMAIN 167 263 IG-LIKE C2-TYPE.  
 FT DOMAIN 31 38 POLY-LEU.  
 FT DOMAIN 289 292 POLY-ALA.  
 FT DISULFID 62 138 POTENTIAL.  
 FT DISULFID 165 243 POTENTIAL.  
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 130 120 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 321 322 HA -> TWAPVYQDYLLPRYLMSPLKTRGLP (in  
 FT isoform 2)  
 FT FTID=VSP\_002521.  
 FT R -> H (IN REF. 4 AND 5; CAC36464).  
 FT CONFLICT 237 237  
 FT SEQUENCE 322 AA; 35960 MW; 55CCBA4AD12E4766 CRC64;  
 SQ SEQUENCE  
 Query Match 7.4%; Score 118; DB 1; Length 322;  
 Best Local Similarity 24.0%; Pred. No. 0.0041;  
 Matches 63; Conservative 33; Mismatches 99; Indels 68; Gaps 11;  
 QY 11 LLLLLLLTRSEVEYRAEYVGNAYLPCGYTPAPAGNL-VFVCGKACPVFEGGNVYL 68  
 DB 35 LLLLSLCAASAEFEVGMVGSNVVLSCLDPHRRHFNLSGLVYV-----QI 81  
 QY 69 RFDENDVNYW-----TSRYLNG-----DRKQDVSLTENVTLADSGICCRQ 113  
 DB 82 ENPEVSVTYTYLPEYKSGINVDSSYNKRGHLSIDSKQGNFSLYLNVVPODQEFCTRV- 140  
 QY 114 IPGINDEKFNKLVYKPAVTPAPTLQRFDPAPRPMLTRGHGPAETQTL-----GSL 168  
 DB 141 ---FMTATELVKILEEYVRLRAVNF-----STPVISTSSNPGORRTTCKSKCY 191  
 QY 169 PDINTQISTANELRDLRLANDLDSGATIRIGYIGAGICAGHALIFGALLPKWYS 228  
 DB 192 PEPNLYWINTDNSLIDTALQNT-----VYLK-----LGLVDVISTLLPMTS 236  
 QY 229 HSK-----EKT-----QNLISLIA 243  
 DB 237 RGDVLCVENVALHQNITISQA 259  
 RESULT 5  
 A33\_HUMAN STANDARD; PRT; 319 AA.  
 AC Q99795;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Cell surface A33 antigen precursor (Glycoprotein A33).  
 GN GPA33.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RP TISSUE=Colon carcinoma;  
 RX MEDLINE=97165045; PubMed=9012807;  
 RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,  
 RA Moritz R.L., Tu G.-F., U.H., Whitehead R.H., Groenen L.C.,  
 RA Scott A.M., Ritter G., Cohen L., Welt S., Old L.J., Nice B.C.,  
 RA Burgess A.W.;  
 RA "The human A33 antigen is a transmembrane glycoprotein and a novel  
 RT member of the immunoglobulin superfamily.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).  
 RN [2]  
 RP POST-TRANSLATIONAL MODIFICATIONS.  
 RX MEDLINE=97396159; PubMed=9245713;  
 RA Ritter G., Cohen L.S., Nice B.C., Catimel B., Burgess A.W.,





DR PDB: 1EAU; 13-JUL-01.  
 DR PDB: 1FSW; 08-NOV-00.  
 DR PDB: 1KAC; 24-NOV-99.  
 DR Genew: HGNC:2559; CXADR.  
 DR MIM: 602621; -.  
 DR GO: GO:0005867; C: integral to plasma membrane; TAS.  
 DR GO: GO:0004872; F: receptor activity; TAS.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003598; IG\_C2.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR Pfam: PF00047; Ig\_2.  
 DR SMART: SM00408; IGC2\_1.  
 DR PROSITE: PS00835; IG\_LIKE; 2.  
 DR Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; signal;  
 Repeat; 3D-structure.  
 FT CHAIN 1 19 POTENTIAL.  
 FT SIGNAL 20 365 EXTRACELLULAR AND ADONOVIRUS RECEPTOR.  
 FT TRANSMEM 238 258 POTENTIAL.  
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 134 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 141 120 IG-LIKE C2-TYPE 2.  
 FT DISULFID 41 120 BY SIMILARITY.  
 FT CARBOHYD 162 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 106 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 365 AA; 40029 MW; AB01C6346CB7FE64 CRC64;

Query Match 7.8%; Score 124.5; DB 1; Length 365;  
 Best Local Similarity 23.7%; Pred. No. 0.0012;  
 Matches 61; Conservative 48; Mismatches 103; Indels 45; Gaps 12;

QY 9 CULLL-----LHLLTRSSSEVRAEYGVNAVLPCEFT--PAAPGNLVPCMGKACGP 59  
 DB 6 CFWLLCGVDFPARSLSTITPEEMIEKAK-GETAYLPCKETLSPEQGLP-DIEMLISPAD 63  
 QY 60 VPECNVLV-RTDERVNVYWTSR---YWLNGDFRKGDSVLTINVTLADSGIYCRR 112  
 DB 64 NQKVQVQVILVSGDKIYDYPDLKGRVHFTSNDLKSQDASINVTNLQLSDIGTYQCKVK 123  
 QY 113 QIRGIMDEKFNKLVKIPKPA---KYTPAPTLQIRP-----TAAFRMLTTRHGPAE 161  
 DB 124 KAPGAN-KKIHVLVLPKSGARCYVDSEETGSPFKICEPEKESLPQYEQMLSDSQ 182  
 QY 162 TQTLGSLDINLTQIS-----TLANELRDSRLANDLRDQATIRIGIYGA 207  
 DB 183 KMPSTWLMEMSSSVSVKASSEVSGTYSCTVRNVGSDGCLRLNVPPSKAGLIGA 242  
 QY 208 GICAGLALALIFGALIF 224  
 DB 243 IIGTLALALILGLIF 258

RESULT 3  
 ID SRBI HUMAN STANDARD; PRT; 398 AA.  
 AC 000241; O8TB12; O9H1U5; O9Y4V0;  
 DT 28-FEB-2003 (Rel. 41; Created)  
 DT 15-SEP-2003 (Rel. 42; Last sequence update)  
 DE 15-SEP-2003 (Rel. 42; Last annotation update)  
 DE Signal-regulatory protein beta-1 precursor (SIRP-beta-1).  
 GN SIRPB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97215901; PubMed=9062191;  
 RA Kharičonenkov A., Chen Z., Sures I., Wang H., Schilling J.,  
 RA Ullrich A.;  
 RA "A family of proteins that inhibit signalling through tyrosine kinase

RT receptors.";  
 RL Nature 386:181-186(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Coty M.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaesialho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.A., Martin S.L., McCormachie L.J., McIay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carinini P., Prange C.,  
 RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP INTERACTION WITH TYROBP AND SYK.  
 RX PubMed=11169422;  
 RA Tomasello E., Cant C., Buehring H.-J., Vely F., Andre P., Seifert M.,  
 RA Ullrich A., Vavler E.;  
 RT "Association of signal-regulatory proteins beta with KARAP/DAP-12.";  
 RL Eur. J. Immunol. 30:2147-2156(2000).  
 RN [5]  
 RP INTERACTION WITH TYROBP, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=20072721; PubMed=10604985;  
 RA Dietrich J., Cella M., Seifert M., Buehring H.-J., Colonna M.;  
 RT "Signal-regulatory protein beta 1 is a DAP12-associated activating  
 RT receptor expressed in myeloid cells.";  
 RL J. Immunol. 164:9-12(2000).  
 CC -I- FUNCTION: Immunoglobulin-like cell surface receptor involved in  
 CC the negative regulation of receptor tyrosine kinase-coupled  
 CC signaling processes. Participates also in the recruitment of



DR SMART; SMO0408; IGC2; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; signal;  
 Repeat.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR  
 FT DOMAIN 20 237 HOMOLOG.  
 FT TRANSMEM 238 258 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 136 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 141 228 IG-LIKE C2-TYPE 2.  
 FT DISULFID 41 120 BY SIMILARITY.  
 FT DISULFID 162 212 BY SIMILARITY.  
 FT CARBOHYD 106 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 201 365 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 340 365 VAPNLSRMGAVPMVIMPAQSKGSIV -> FYAVKTDGDT  
 FT (IN REF. 2 AND 3).  
 SQ SEQUENCE 365 AA; 39947 MW; 5445B4B52A34B2A2 CRC64;  
 Query Match 8.3%; Score 131.5; DB 1; Length 365;  
 Best Local Similarity 23.3%; Pred. No. 0.00029;  
 Matches 65; Conservative 47; Mismatches 96; Indels 71; Gaps 14;  
 QY 9 CYLLLL-----LLLTRSEVEYRAEYGNATLPCTT--PAAPNLUVVCWKGAC 58  
 DB 6 CVELLCGIADFTSGISITTEQRIE-KAK-GETAYLPCKFTLSPDOGL-DIEW----- 57  
 QY 59 PVFECCGNYVLTREDDV-----NYMTSR-----YMLNGDPFKGVSLTIENV 100  
 DB 58 -----LISPDNDQIVQVILYSGDKTYNDYEDLKGKRVFTSNDVKSAGASINVTLL 110  
 QY 101 TLADSGIYCCRI-QIPGIMNDEKFNKLVKPAK-----VTPAPTLQDF-----TAAF 148  
 DB 111 QLSIDIGTQCKYKAPGVAN-KKFLTLTVLVKPSGRFCFVDSGEIEGNDPKLKCEPREGSL 169  
 QY 149 PPMULTTRHGSPAEFTQTLGSLPFINLTQIS-----TLANEIRDRLANDLRD 194  
 DB 170 PLQFQWQKLSQSTQPTPWLAMETSPVISVKNASSEYSGTYSCTQVNRVGSQCMRLRDV 229  
 QY 195 SGATIRIGIYIGAGICAGLALALIFGALIFKMYSHSKK 233  
 DB 230 VPPSNRAGTLAGAVIGTLALVLI-GALIFCCHRRKREE 267  
 RESULT 2  
 CXAR\_HUMAN STANDARD; PRT; 365 AA.  
 ID P78310; O00694;  
 AC P78310; O00694; Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Coxsackievirus and adenovirus receptor precursor (Coxsackievirus B-  
 adenovirus receptor) (hCAR) (CVB3 binding protein).  
 GN CXAR OR CAR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97190109; PubMed=9036860;  
 RA Beigelson J.M., Cunningham J.A., Drogue G., Kurt-Jones B.,  
 RA Kitchin J.S., Hong J.S., Horwitz M.S., Crowell R.L., Flanberg R.W.;  
 RT "Isolation of a common receptor for Coxsackie B viruses and  
 RT adenoviruses 2 and 5.";  
 RT Science 275:1320-1323 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97250541; PubMed=9096397;  
 RA Tomko R.P., Xu R., Phillips L.;  
 RT "hCAR and hCAR: the human and mouse cellular receptors for subgroup C  
 RT adenoviruses and group B coxsackieviruses.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20008750; PubMed=10543405;  
 RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,  
 RA Bowles N.E.;  
 RT "Genomic organization and chromosomal localization of the human  
 RT Coxsackievirus B-adenovirus receptor gene.";  
 RT Hum. Genet. 105:354-359 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX Anderson C.W., Kielczawa J., Dunn J.J., Freimuth P.;  
 RT "Sequence and expression of CXAR, the human gene for the  
 RT coxsackievirus and adenovirus receptor.";  
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX Anderson B., Tomko R., Anderson K., Darban H., Oncu D., Mizra M.,  
 RA Solterprant K., Somhammer E., Phillips L.;  
 RT "Putative regulatory domains in the human and mouse CAR genes.";  
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Caranci P., Prange C.,  
 RA Raha S.S., Loguett N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinchman J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND  
 CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; Y07593; CAA6868.1; -;  
 DR EMBL; U90716; AAC51234.1; -;  
 DR EMBL; AF169366; AAF05908.1; -;  
 DR EMBL; AF169360; AAF05908.1; JOINED.  
 DR EMBL; AF169361; AAF05908.1; JOINED.  
 DR EMBL; AF169362; AAF05908.1; JOINED.  
 DR EMBL; AF169363; AAF05908.1; JOINED.  
 DR EMBL; AF169364; AAF05908.1; JOINED.  
 DR EMBL; AF169365; AAF05908.1; JOINED.  
 DR EMBL; AF200465; AAF24344.1; -;  
 DR EMBL; AF242865; AAG01088.1; -;  
 DR EMBL; AF242862; AAG01088.1; JOINED.  
 DR EMBL; AF242864; AAG01088.1; JOINED.  
 DR EMBL; BC003684; AAH03684.1; -;  
 DR EMBL; BC010536; AAH10536.1; -;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 01:50:35 ; Search time 14.4811 Seconds  
(without alignments)  
977.484 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587  
Sequence: 1 MSHLPDCCVLLLLLLTR.....CVSSRQPSQPLGCRFAMP 301

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131.5	8.3	365	1	CGAR_MOUSE
2	124.5	7.8	365	1	CGAR_MOUSE
3	118.5	7.5	398	1	SRL1_HUMAN
4	118	7.4	322	1	ICOL_MOUSE
5	112.5	7.1	319	1	A33_HUMAN
6	109	6.9	530	1	PVR2_MOUSE
7	108	6.8	1138	1	TIR1_HUMAN
8	107	6.7	310	1	HEMA_RACV1
9	102.5	6.5	555	1	C166_CARAV
10	101.5	6.4	538	1	PVR2_HUMAN
11	100.5	6.3	387	1	SRL2_HUMAN
12	99	6.2	818	1	TRKB_CHICK
13	98.5	6.2	245	1	MOG_RAT
14	98	6.2	526	1	BUTY_HUMAN
15	97.5	6.1	315	1	HEMA_VACCT
16	97	6.1	506	1	SHS1_BOVIN
17	95	6.0	467	1	SIL5_MOUSE
18	95	6.0	503	1	SHS1_HUMAN
19	94.5	6.0	246	1	MOG_MOUSE
20	94.5	6.0	773	1	PIGR_RABIT
21	94	5.9	215	1	CIB2_RAT
22	94	5.9	246	1	MOG_BOVIN
23	93.5	5.9	349	1	CEA8_HUMAN
24	93	5.9	313	1	HEMA_VARV
25	92.5	5.8	306	1	CD80_MOUSE
26	92	5.8	1010	1	CONT_CHICK
27	92	5.8	1134	1	TIR1_MOUSE
28	91.5	5.8	459	1	CD4_RABIT
29	91	5.7	344	1	CEA6_HUMAN
30	90.5	5.7	519	1	ECTO_RAT
31	89.5	5.6	302	1	ICOL_HUMAN
32	89.5	5.6	314	1	HEMA_VACCV
33	89.5	5.6	344	1	NTRI_HUMAN

34	89	5.6	564	1	C166_BRARE	Q90460 brachydanio
35	88.5	5.6	344	1	NTRI_RAT	Q62718 ratu mus musculu
36	88.5	5.6	771	1	PIGR_MOUSE	O60939 homo sapien
37	88	5.5	215	1	CIB2_HUMAN	P20978 vaccinia vi
38	88	5.5	315	1	HEMA_VACCT	Q951h0 pan troglod
39	88	5.5	597	1	SIL5_PANTR	Q99pj0 mus musculu
40	87.5	5.5	344	1	NTRI_MOUSE	O63994 mus musculu
41	87.5	5.5	403	1	CD33_MOUSE	P13688 homo sapien
42	87.5	5.5	526	1	CEA1_HUMAN	P08714 vaccinia vi
43	86.5	5.5	315	1	HEMA_VACCV	O61490 mus musculu
44	86.5	5.5	583	1	C166_MOUSE	O60469 homo sapien
45	86	5.4	2012	1	DSCA_HUMAN	

## ALIGNMENTS

```
RESULT 1
CXAR_MOUSE          STANDARD;          PRT;          365 AA.
ID AC P97752; 009052;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coxsackievirus and adenovirus receptor homolog precursor (mCAR).
GN CXADR OR CAR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=97190109; PubMed=9036860;
RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,
RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;
RT "Isolation of a common receptor for Coxsackie B viruses and
RT adenoviruses 2 and 5."
RT Science 275:1320-1323(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/MAT;
RX MEDLINE=97250541; PubMed=9096397;
RA Tomko R.P., Xu R., Philipson L.;
RT "HCAR and mCAR: the human and mouse cellular receptors for subgroup C
RT adenoviruses and group B coxsackieviruses."
RT Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA Bergelson J.M., Krithivas A., Crowell T.L., Finberg R.W.;
RT "The murine CAR homologue (mCAR) is a receptor for coxsackie B
RT viruses and adenoviruses."
RT Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Y10320; CAA71368.1; -
CC EMBL; U90715; AAC53148.1; -
CC EMBL; Y11929; CAA72679.1; -
CC MGD; MGI:1201679; Cxadr.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 2.
```

A: Molecule type: mRNA  
A: Residues: 1-315 <JIN>

A: Cross-references: GB:X15709; GB:M5773; NID:G61313; PID:CAA33740.1; PID:G61314

C: Superfamily: vaccinia virus hemagglutinin; immunoglobulin homology

C: Keywords: glycoprotein; hemagglutinin; late protein; transmembrane protein

F: 1-16/Domain: signal sequence #status predicted <SIG>

F: 17-315/Product: hemagglutinin #status predicted <MAT>

F: 27-105/Domain: immunoglobulin homology <IMM>

F: 280-302/Domain: transmembrane #status predicted <TM>

F: 303-315/Domain: intracellular #status predicted <INT>

F: 37,69,112,161,254/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.1%; Score 97.5; DB 1; Length 315;

Best Local Similarity 17.4%; Pred. No. 1.1; Mismatches 116; Indels 97; Gaps 12;

Matches 57; Conservative 57; Mismatches 116; Indels 97; Gaps 12;

3 SHLPDCVLLLLLTSSSEVEYRAEYVGNAYLPCEFYTPAPAGN-----LVPCWCK 55

2 ARLP---ILLILSLVSTPSPQTSKIDDLSC-----NRNNINDYVWMSMYK 50

56 GACPVFECGNVLTDERDV---NYMTRWYLANGDFRKGVSLTIENVTLADSGIYCCR 111

51 -----EPNSITLLAKSDVLFEDNYTKKISYDSPYDDLVTITIKSLTARDAGYVCA 104

112 IOIPGIMND-----EKFKLVIVIPAKVTAPFTLQRFDAFPFRLTLTRGHPAETQL 165

105 FPMTPPTDTRKVDYEEYESTELIV-----NTDSESTIDILSGSTHSP---ETS 150

166 GSLP-DINLTQISTLANELRDSRLANDSGATIRIGIYGAGICAGLALIFGALIF 224

151 SEKPEDIDVNCSSVFEITPPTIDNVEDHTDV----- 185

225 KMYSHSKEKIQNLISLISLANLPPSGLANAVAGIRSEENIYTIENV----- 271

186 ---TYSDSINTVSATS-----GESTDETPEPTDKEDHTVTPVSYYTSTSSGIYT 237

272 --YVEEPEYECYVSSRQPSQPLGC 296

238 TKSTDDADLYDPTVNDNDTPSTTVGC 264

#### RESULT 15

T28598

hypothetical protein J7R - variola major virus

C: Species: variola major virus

C: Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #exc\_change 21-Jul-2000

C: Accession: T28598

R: Masung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin

A: Title: Potential virulence determinants in terminal regions of variola smallpox virus

A: Reference number: Z20488; PMID:94088747; PMID:8264798

A: Accession: T28598

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Cross-references: EMBL:L22579; NID:G623595; PID:AAA60908.1; PID:G439077

A: Experimental source: strain Bangladesh 1975

C: Superfamily: vaccinia virus hemagglutinin; immunoglobulin homology

Query Match 6.0%; Score 96; DB 2; Length 313;

Best Local Similarity 21.5%; Pred. No. 1.4; Mismatches 62; Indels 42; Gaps 8;

Matches 38; Conservative 35; Mismatches 62; Indels 42; Gaps 8;

11 LLLLLLTSS---EVEYRAEYVGNAYLPCEFYTPAPAGN---VPVCMGKACPVFEC 63

6 ILLILSLVSTPPIQISKIGDDATLSC-----SRNNINDYVWMSMYK-----EP 54

64 GNVVLTDERDV---NYMTRWYLANGDFRKGVSLTIENVTLADSGIYCCRIGIPIGIMN 119

55 NSIILAAKSDVLYFPDNYTKKISYDSPYDDLVTITIKSLTARDAGYVCAFFMTSTN 114

120 D-----EKFKLVIVIPAKVTAPFTLQRFDAFPFRLTLTRGHPAETQLGLSPD 170

DB 115 DTDKVDYEEYESTELIV-----NTDSESTIDILSGSSHSP---ETSSKPD 157

Search completed: November 22, 2003, 05:44:26  
Job time: 27.8247 secs

A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-818 <V12>  
A:Cross-references: EMBL:X77251; NID:9563881; PIDN:CAA54468.1; PID:g472934  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994  
R:Dechant, G.; Biffo, S.; Okazawa, H.; Kolbeck, R.; Potgiesser, J.; Barde, Y.A.  
Development 119, 545-558, 1993  
A>Title: Expression and binding characteristics of the BDNF receptor chick trkB.  
A:Reference number: S42175; MUID:94116452; PMID:8287802  
A:Accession: S42175  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-818 <DEC>  
A:Cross-references: EMBL:X74109; NID:g407798; PIDN:CAA52210.1; PID:g407799  
A:Gene: trkB  
A:Function:  
A:Description: regulation of nervous system development; receptor for brain-derived neurotrophic factor  
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein  
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor kinase  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-818/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>  
F:32-432/Domain: extracellular #status predicted <EXT>  
F:66-90/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
F:91-114/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
F:115-137/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
F:138-149/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR4>  
F:433-449/Domain: transmembrane #status predicted <TM>  
F:450-818/Domain: cytosolic #status predicted <CT>  
F:532-810/Domain: protein kinase homology <KIN>  
F:540-548/Region: protein kinase ATP-binding motif  
F:566,94,120,199,204,253,287,324,337,349,408/Binding site: carbohydrate (Asn) (covalent)  
F:566/Active site: lys #status predicted  
F:702/Binding site: phosphate (Tyr) (covalent) #status predicted  
F:813/Binding site: phosphate (Tyr) (covalent) #status predicted  
Query Match 6.2%; Score 99; DB 1; Length 818;  
Best Local Similarity 19.9%; Pred. No. 2.6;  
Matches 59; Conservative 41; Mismatches 89; Indels 108; Gaps 12;  
QY 87 DPRKGDVLTENNVLADSGIT--CCRIQIPG-INNDEK-----NLKLVTKP----- 131  
DB 241 DTSKNPASHLTKNVSMSMGWISCAENIVGEVOTSALTVFAPNITFIESPDPDHW 300  
QY 132 -----AKTPAPTLGRDFTAPR-----M 151  
DB 301 CIPFVKGKPKFTLQWTFBGLINSEYITCKIHVINOSEYHGCLOLDPHNLNGAYTL 360  
QY 152 LTRGHGPAETQT-----LGSPLDINI/IOISTLANELRDSR-----LAN 190  
DB 361 LAKNEYGEDEKRVADHFMSPVPGDSGPVDPVVEYETTPNDLGDITNNSNGITSPDVSN 420  
QY 191 DLKDSGATIRIGIYAGAGICAGLALALIFGALIFKMYSHSKKIKONLSIT----- 240  
DB 421 KENEDSTIVYVVGAIALVCTGLVITMLI-----ILKFRGRSHKMGKMPSPSVISNDDSDASP 476  
QY 241 -----SLANLPPS--GLANVAEGIRSEENITYIBENYVEPEPEYCVYSSROOP 290  
DB 477 LHHISNGSTBSSSEGGPDVITIGM-----TKIPVLENP-QYGEITNSQKLP 522

## RESULT 12

161783

sodium channel beta 2 subunit - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Nov-1999

C:Accession: I61783

R:Issom, L.L.; Ragsdale, D.S.; DeLong, K.S.; Westendorp, R.E.; Reber, B.F.X.; Scheuer, A.; Title: Structure and Function of the Beta 2 Subunit of Brain Sodium Channels, a Transmembrane Protein; MUID:96067641; PMID:8521473

A:Accession: I61783

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-186 <RES>  
A:Cross-references: EMBL:U37147; NID:g1086498; PIDN:AA60506.1; PID:g1086499  
A:Gene: SCN2  
A:Introns: 50/3; 120/2

Query Match 6.2%; Score 98; DB 2; Length 186;  
Best Local Similarity 22.9%; Pred. No. 0.5;  
Matches 50; Conservative 21; Mismatches 69; Indels 78; Gaps 9;

QY 31 GQNAVLPCTYPAPAGN--LVPVCWKGACAPFECGN-----VLTDRDENVYVTSRYW- 83  
DB 14 GSDTRLPCTFMSCVVNHKQFSIWN-----TYQESNCSSEENVLQFRKIKNLKLERFGD 68  
QY 84 ---LNGDRKGDVLTENNVLADSGITCCRIQIPGINNDEKFKLVKPAKTPAPTL 140  
DB 69 RVFESGNSKYDVSTLNVQLEDEGINCYITNP-----PDR----- 106  
QY 141 QRDFAAPRMLTTRGHGPAETQTGSLPDINL/IOISTLANELRDSRLANDLDSGATIR 200  
DB 107 -----HRGHKITYLVLEVP-----PERDSTVAVT 132  
QY 201 IGIVYAGICAGLALALIFGALIFKMYSHSKKIKONLS 238  
DB 133 VGASV-----GFLAVVILVLMVVCVARKE--QKLS 163

## RESULT 13

S70587

butyrophilin precursor - human

C:Species: Homo sapiens (man)

C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 21-Jul-2000

C:Accession: S70587

R:Taylor, M.R.; Peterson, J.A.; Ceriani, R.L.; Couto, J.R.  
Biochim. Biophys. Acta 1306, 1-4, 1996

A&gt;Title: Cloning and sequence analysis of human butyrophilin reveals a potential receptor

A:Reference number: S70587; MUID:96201696; PMID:8611614

A:Accession: S70587

A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-526 <TAY>  
A:Cross-references: EMBL:U39576; NID:g1326082; PIDN:AAC50489.1; PID:g1326083

Query Match 6.2%; Score 98; DB 2; Length 526;  
Best Local Similarity 28.4%; Pred. No. 1.8;  
Matches 40; Conservative 21; Mismatches 58; Indels 22; Gaps 6;

QY 3 SHLPDVCYLILLLILTRSSSEY-----RAVGNAYLPCFYTPAPAGNLVPCW 53  
DB 7 SGLP-RCLTLTILIQPLTDSAPDPVIGPPPIYALVGEDLELPRSLPMSAEHLRLW 65  
QY 54 -GKACPVFEGCGVNLRTDERD-----VNYTSRYWLNQD-FRKGDVLTENNVLADSG 106  
DB 66 FRKVVSPA-----VLVHRDGRQEAQMPFVRGATLVQDGIAGKRVALLRIGVAVSDG 120  
QY 107 IYCCRIQIPGINNDEKFKLV 127  
DB 121 EYTCFREDGSYEALVHLKV 141

## RESULT 14

HN42VT

hemagglutinin precursor - vaccinia virus (strain Tianan)

C:Species: Vaccinia virus

C:Date: 30-Jun-1990 #sequence\_revision 30-Jun-1990 #text\_change 16-Jul-1999

C:Accession: J01018

R:Lin, D.; Li, Z.; Jin, O.; Yuwen, H.; Hou, Y.  
J. Exp. Med. 170, 571-576, 1989A>Title: Vaccinia virus hemagglutinin. A novel member of the immunoglobulin superfamily.  
A:Reference number: J01018; MUID:89328331; PMID:2754392

A:Accession: J01018

Db 364 KRRSDGKHKVLTCAESGPKPDVMSVNGTNDVSYNNKA---TYKLTIVPSKNTLVS 440  
Qy 177 STLANEL-----RDSRLANDLRDGS---ATIRIGIYGAGICAGLALIFGALIFK---- 225  
Db 441 CLVTKLKGBDTRKEISVFSQKNEDEGEQAKVIVIGVLLVAAAL-VGIIYWIYIKKTGQ 499  
Qy 226 -WYSHSK 232  
Db 500 SWKTGEKE 507

RESULT 8  
153960  
P:PRR2 alpha - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jan-2000  
C:Accession: 153960  
R:Eberle, F.; Dubreuil, P.; Mattei, M.G.; Devillard, E.; Lopez, M.  
Gene 159, 267-272, 1995  
A>Title: The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is ch  
A:Reference number: 153960; MUID:95347610; PMID:7622062  
A:Accession: 153960  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-478 <RES>  
A:Cross-references: GB:S79171; NID:G1042202; PID:G1042203  
C:Genetics:  
A:Gene: PRR2alpha  
C:Superfamily: poliovirus receptor; immunoglobulin homology  
F:276-331/Domain: immunoglobulin homology <IMM>

Query Match 6.4%; Score 101.5; DB 2; Length 478;  
Best Local Similarity 25.6%; Pred. No. 0.81;  
Matches 57; Conservative 29; Mismatches 96; Indels 41; Gaps 13;  
Qy 11 LLLLLLLTRSEV-----EYRAVGQNAVLPCFYTPAAGNLVP-VCMGKACPVFEC 63  
Db 21 LLLLLLLETGADVRVQVLPVVRGQGGTVLPCHLPLPVGLVLSVLTWQRPDAPA-NH 79  
Qy 64 GNVVL-----RTDERDVNYWTSRYWLNQDFRK--GDVSLTIENVTLADSGIYC 109  
Db 80 QNVAAFHFKMGSPSPKSGSERLSFVSAKOSTGQDTAELODATALHGLTVEDEGNYT 139  
Qy 110 CRI-----QIPGINDEKFNKLKVIKPAKTPAP--TLQDFTAAFPRLMTRGHGPAE 161  
Db 140 CEFATFPKGSVRGWT-----WLRVIKPKQAQAEQKVFSPGDPPTV-ALCISKEGPPAR 193  
Qy 162 TOTGLSLP-DINLTQIS-TLANELR-DSR--LANDLRDSCATI 199  
Db 194 ISWLSLDMWEAKETQVSGTLGTVVTSRFTLVPSGRADGTV 236

RESULT 9  
168093  
P:PRR2 delta - human  
C:Species: Homo sapiens (man)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jan-2000  
C:Accession: 168093  
R:Eberle, F.; Dubreuil, P.; Mattei, M.G.; Devillard, E.; Lopez, M.  
Gene 159, 267-272, 1995  
A>Title: The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is ch  
A:Reference number: 153960; MUID:95347610; PMID:7622062  
A:Accession: 168093  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-538 <RES>  
A:Cross-references: GB:S79172; NID:G1042204; PID:G1042205  
C:Genetics:  
A:Gene: PRR2delta  
C:Superfamily: poliovirus receptor; immunoglobulin homology  
F:276-331/Domain: immunoglobulin homology <IMM>

Query Match 6.4%; Score 101.5; DB 2; Length 538;  
Best Local Similarity 25.6%; Pred. No. 0.93;  
Matches 57; Conservative 29; Mismatches 96; Indels 41; Gaps 13;  
Qy 11 LLLLLLLTRSEV-----EYRAVGQNAVLPCFYTPAAGNLVP-VCMGKACPVFEC 63  
Db 21 LLLLLLLETGADVRVQVLPVVRGQGGTVLPCHLPLPVGLVLSVLTWQRPDAPA-NH 79  
Qy 64 GNVVL-----RTDERDVNYWTSRYWLNQDFRK--GDVSLTIENVTLADSGIYC 109  
Db 80 QNVAAFHFKMGSPSPKSGSERLSFVSAKOSTGQDTAELODATALHGLTVEDEGNYT 139  
Qy 110 CRI-----QIPGINDEKFNKLKVIKPAKTPAP--TLQDFTAAFPRLMTRGHGPAE 161  
Db 140 CEFATFPKGSVRGWT-----WLRVIKPKQAQAEQKVFSPGDPPTV-ALCISKEGPPAR 193  
Qy 162 TOTGLSLP-DINLTQIS-TLANELR-DSR--LANDLRDSCATI 199  
Db 194 ISWLSLDMWEAKETQVSGTLGTVVTSRFTLVPSGRADGTV 236

RESULT 10  
JC2457  
vascular cell adhesion protein - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 15-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 08-Oct-1999  
C:Accession: JC2457  
R:Tsang, Y.T.M.; Haskard, D.O.; Robinson, M.K.  
Biochem. Biophys. Res. Commun. 201, 805-812, 1994  
A>Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule.  
A:Reference number: JC2457; MUID:94271236; PMID:7516159  
A:Accession: JC2457  
A:Molecule type: mRNA  
A:Residues: 1-538 <TSA>  
A:Cross-references: EMBL,U08351; NID:G474382; PIDN:AAA21542.1; PID:G474383  
C:Keywords: glycoprotein; transmembrane protein  
F:497-517/Domain: transmembrane #stratus predicted <TM>  
F:75,157,271,330,360/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 101; DB 2; Length 538;  
Best Local Similarity 25.6%; Pred. No. 1;  
Matches 61; Conservative 29; Mismatches 78; Indels 70; Gaps 13;  
Qy 85 NGD--FRKGDVSLTIENVTLADSGIYCCRIQIPGINDEKFNKLKVIKPAKTPAPTQOR 142  
Db 262 NGDQQLLGGNATLTLLIARMEDSGIYCEGVNPGVGNKREVELTVQVAP-----R 311  
Qy 143 DETAAPRLMTRGHGPAETQTLGSLPDINTQIS-----TLANELRDSRL----- 188  
Db 312 D-----TIIISVPSSTLEBGS--SVNMTCSGDFPAPKILMSKLRDGNLPLESEN 360  
Qy 189 -----ANDLRDSCATIRIGIYGAGI-CAGL--ALALIFGA---LIFKWSHSEKI 234  
Db 361 TTLTLTSTRKEDS-----GIYCEGINQAQINRKEVELLIQAAPKDLQTLTFPESVYE 414  
Qy 235 QNLSLIS--LANLPSGL-----ANVAAGIRSEENITYT-----EENVYVEEPPN 279  
Db 415 GDTVIISCTCGVPEPTLLILKKKATGDTVKSTGAVTIHRARLADGAVVECESKNE 472

RESULT 11  
S44098  
brain-derived neurotrophic factor receptor precursor - chicken  
N:Alternate names: receptor tyrosine kinase trkB  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Gallus gallus (chicken)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 11-Jun-1999  
C:Accession: S59938; S42175; S44098  
R:Vinik, N.O.; Erdmann, K.S.; Heumann, R.  
Gene 149, 383-384, 1994  
A>Title: Cloning and sequence analysis of a cDNA encoding a novel truncated form of the  
A:Reference number: S59938; MUID:95047511; PMID:7959025  
A:Accession: S59939

protein-tyrosine kinase (EC 2.7.1.112), receptor type tie precursor - human

C.Species: Homo sapiens (man)

C.Date: 19-Feb-1994 #sequence revision 11-Aug-1995 #text\_change 21-Jul-2000

C.Accession: S24066; C38269; I52613

R.Paratamen, J.; Armstrong, E.; Maekelae, T.P.; Korhonen, J.; Sandberg, M.; Renkonen, R.; Mol. Cell. Biol. 12, 1698-1707, 1992

A.Title: A novel endothelial cell surface receptor tyrosine kinase with extracellular ep

A.Reference number: S24066; MUID:92195316; PMID:1312667

A.Accession: S24066

A.Molecule type: mRNA

A.Residues: 1-1138 <PAR>

A.Cross-references: EMBL:X60957

R.Paratamen, J.; Maekelae, T.P.; Aitalo, R.; Lehtvaeslahti, H.; Aitalo, K. Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990

A.Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.

A.Reference number: A38268; MUID:91062389; PMID:2247464

A.Accession: C38269

A.Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A.Molecule type: mRNA

A.Residues: 981-1034 <PAM>

A.Experimental source: clone JTK14

R.Korhonen, J.; Lahtinen, I.; Halmekyto, M.; Alhonen, L.; Janne, J.; Dumont, D.; Aitalo

Blood 86, 1828-1835, 1995

A.Title: Endothelial-specific gene expression directed by the tie gene promoter in vivo

A.Reference number: I52613; MUID:95383653; PMID:7655012

A.Accession: I52613

A.Status: translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-19 <RFS>

A.Cross-references: GB:S79347; NID:g1086921; PIDN:ADD14299.1; PID:g4261999

C.Genetics:

A.Gene: GDB:TIE; JTK14

A.Cross-references: GDB:212873; OMIM:600222

A.Map position: 1p34-1p33

C.Function:

A.Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP

C.Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin tyro

C.Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph

F.1-21/Domain: signal sequence #status predicted <SIG>

F.32-1138/Product: protein-tyrosine kinase, receptor-type tie #status predicted <MOT>

F.126-107/Domain: immunoglobulin homology <IM1>

F.215-255/Domain: EGF homology <EG1>

F.259-302/Domain: EGF homology <EG2>

F.306-344/Domain: EGF homology <EG3>

F.365-428/Domain: immunoglobulin homology <IM2>

F.449-530/Domain: fibronectin type III repeat homology <FN3>

F.542-633/Domain: fibronectin type III repeat homology <FN2>

F.642-730/Domain: fibronectin type III repeat homology <FN3>

F.761-786/Domain: transmembrane #status predicted <TM>

F.837-1114/Domain: protein kinase homology <KIN>

F.845-853/Region: protein kinase ATP-binding motif

F.443-105.372-426/Diulfide bonds: #status predicted

F.83.161.503.596.709/Binding site: carbohydrate (Asn) (covalent) #status predicted

F.870.887.979/Active site: Lys, Glu, Asp #status predicted

Query Match 6.8%; Score 108; DB 1; Length 1138;

Best Local Similarity 23.6%; Pred. No. 0.65;

Matches 47; Conservative 32; Mismatches 74; Indels 46; Gaps 8;

17 LTRTSSEVEYRAEYGNAYLPCFYTPAAPGLVPYCWKGACPFVECGNVVLRTERDYN 76

135 ILMNASELEFNLETPMR--INC-----AAAGNPFPV---RGSIELRKPDPGVLLSTKAVI- 402

77 YMYSRWINDPFKRGDVSRTIENTLADSGIYCCRCIQIPIGIMNDEKFNILKVIKPKATVP 136

403 -----EPEPTTAAFEFVPRVLVADSGFWECKVSTSGGDSRRFRKVV-----KVP 447

137 APTLQRFDTAPFRLMTTR-----GHGPAETQTLGSLPDIVLQISTLANEL 183

448 VP-----LAARLLRLKQSRQLVSPVSVSGDGPISITVRLHYRPQDSITMWSITIVDP 500

184 RDSRLANDIR-DSGATIRI 201

Db 501 SENVTLMTLRPKTGYSVRV 519

RESULT 6  
HNVZRA  
hemagglutinin precursor - raccoonpox virus  
C|Species: raccoonpox virus  
C|Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 16-Jul-1999  
C|Accession: A43381  
R|Cavallaro, K.F.; Esposito, J.J.  
Vi.tology 190, 434-439, 1992  
A|Title: Sequences of the raccoon poxvirus hemagglutinin protein.  
A|Reference number: A43381; MUID:92410621; PMID:1529542  
A|Accession: A43381  
A|Molecule type: DNA  
A|Residues: 1-310 <CAV>  
A|Cross-references: GB:M94169; NID:g333625; PIDN:AAA47231.1; PID:g333626  
C|Superfamily: vaccinia virus hemagglutinin; immunoglobulin homology  
C|Keywords: glycoprotein; hemagglutinin; late protein; transmembrane protein  
F|1-16/Domains: signal sequence #status predicted <SIG>  
F|17-310/Product: hemagglutinin #status predicted <MAT>  
F|29-107/Domains: immunoglobulin homology <IMM>  
F|327-293/Domains: transmembrane #status predicted <TM>  
F|36-105/Dissulfide bonds: #status predicted  
F|39,113,133,203/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7%; Score 107; DB 1; Length 310;  
Best Local Similarity 23.2%; Pred. No. 0.16;  
Matches 53; Conservative 36; Mismatches 81; Indels 58; Gaps 10;

QY 10 VLLILLILITRSSEVERAEVGONAVLPCTTPAAPGNLVPCWKGACPVFECCNVLR 69  
DB 8 ILLLSIVTTTKPHPTQISKLGDAITLSC-NRNTHGYLVWSWYK-----KPDSIILL 60  
QY 70 TDERDVNWTSRWMLNGDFPKGVS-----LTIEVNTLADSGICRCIQIPGLM 118  
DB 61 AAANDVVYF-----DDYADRKYSDSPYDTLATITITSLTSADACTYVCAPFTS-T 112  
QY 119 ND-----EKFNKLVIKPAKV-----PAPTQRDFTAFPMULTTRGHGAET-QTL 165  
DB 113 NOTDKIDVEYFDVLVNPNVSTIDAILSGSTQQOIIISH-----TEQHDSPTTICTS 167  
QY 166 GSLPDINLTQISTLANELRDS-----RLANDIRDSGATT 199  
DB 168 ESTTQISESTSETSSQISSETSESYGVEDRDQYNVTPTTDNSDTI 215

RESULT 7  
150478  
neuroilin - goldfish (fragment)  
C|Species: Carassius auratus (goldfish)  
C|Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1996  
C|Accession: I50478  
R|laessing, U.; Giordano, S.; Stecher, B.; Lotspeich, F.; Stuermer, C.A.  
Differentiation 56, 21-29, 1994  
A|Title: Molecular characterization of fish neuroilin: a growth-associated cell surface protein DM-GASP/SC-1/BEN.  
protein DM-GASP/SC-1/BEN.  
A|Reference number: I50478; MUID:94299040; PMID:8026643  
A|Accession: I50478  
A|Status: preliminary; translated from GB/EMBL/DDBJ  
A|Molecule type: mRNA  
A|Residues: 1-523 <LAE>  
A|Cross-references: GB:I25056; NID:g407318; PID:g407319

Query Match 6.5%; Score 102.5; DB 2; Length 523;  
Best Local Similarity 23.4%; Pred. No. 0.74;  
Matches 44; Conservative 26; Mismatches 75; Indels 43; Gaps 7;

QY 78 WTSRWMLNDPRKGVSLTIENVTLADSGIYCRIQIPGINDEKENL----- 125  
DB 330 WK-----DNKRDLDPFSKITYS DAGLYVDVSI EGIKRSLSFELTVGCIPIKITSYT 383  
QY 126 -----KLVIKPAKVTPAPTLQRDFTAFPMULTTRGHGAETQTIGSLPDINTQI 176

```

QY      247  P S G L A N A V A E G I R S E N I Y T I E E N Y E E E   276
          | | | | : : : | | | | : : :
Db      412  F K T L Q N A V K K E V H A E D N I Y - I E N N L A M N Q   440

```

## RESULT 2

Coxsackie - and adenovirus receptor - bovine  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 02-Apr-2002  
C/Accession: J07780  
R/Inoeden, I.; Keyaerts, E.; Lindberg, M.; Van Ranst, M.  
Biochem. Biophys. Res. Commun. 288, 805-808, 2001  
A>Title: Characterization of a cDNA encoding the bovine coxsackie and adenovirus recept  
A/Reference number: J07780

A:Residues: 1-365 <THO>  
A:Cross-references: GB:AY033651  
C:Comment: This protein serves as the primary adenoviral attachment site on bovine cells

Query Match	7.6%	Score 120	DB 2	Length 365
Best Local Similarity	23.9%	Pred. No.	0.015	
Matches 64	Conservative 44	Mismatches 110	Indels 50	Gaps 13

0Y 11 LLLLLL-----TRSSSEVERAEV-----GQNAVLPCTYT--PAAPGNLVPCWCKGA 57

Db 3 LLLRFLLCGVADFTRGSLITTPQMIKAKGETAYLPCKFTLGPEDQGPL-DIEWLSP 61

58 CPVFECGNVL--RTDERDVNYWTSR----YWLNGDFRKGDVSLTIENVTLADSGIYCCR 11

db 62 ADNQKVDQVILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCK 12

QY 112 I-QIPGIMNDEKFNKLVIKPA---KVTPTATLQRF-----TAAFPRLTTRGHCP 15

Db 122 VKAPGVGN-KKIQLTVLKP SGIRCYVDGSEEIGNDFKLKCEPKESLPLRYEQKSD 18

QY 160 AETQTLGSLPDINLTQIS-----TLANELRDSRLANDLRDSGATIRIGIYI 20

Db 181 SOKLPTSWLEPMTSPVISVKNASAEYSGTYTCTVRNRVGSDOCLRLDVPPSNRAGTIA 24

206 GAGCAGLALATIFGALIFKWSHSEK 233

Db 241 GAVT--GTTATATATV-ECCHKKRR 265

### RESULT 3

poliovirus receptor homolog precursor ( mouse  
 C:Species: Mus musculus domesticus (western European house mouse)  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 22-Jun-1999  
 C:Accession: A38211  
 R:Morrisson, M.E.; Racanelli, V.R.  
 J. Virol. 66, 2807-2813, 1992  
 A:Title: Molecular cloning and expression of a murine homolog of the human poliovirus re  
 A:Reference number: A38211; MUID:92219365; PMID:1560525  
 A:Accession: A38211  
 A:Molecule type: DNA  
 A:Residues: 1-467 <MOR>  
 A:Cross-references: GB:M80206; NID:G199785; PID:AAA39734.1; PID:G199786  
 C:Superfamily: poliovirus receptor; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-467/Product: poliovirus receptor homolog #status predicted <MAT>  
 F:26-354/Domain: extracellular #status predicted <EXT>  
 F:47-133/Domain: immunoglobulin homology <IMM1>  
 F:167-231/Domain: immunoglobulin homology <IMM2>  
 F:267-332/Domain: immunoglobulin homology <IMM3>  
 F:355-374/Domain: transmembrane #status predicted <TMN>  
 F:375-467/Domain: intracellular #status predicted <INT>  
 F:554-131, 174-229, 274-320/Disulfide bonds: #status predicted  
 F:120, 138, 315/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match	6.9%;	Score 109;	DB 1;	length 467;
Best Local Similarity	22.0%;	Pred. No. 0.18;		
Matches	63;	Conservative	33;	Mismatches 103;
				Indels 88;
				Gaps 10;

QY 11 LLLLLLLTRSSSEV-----EYRAEVGQNAYLPCFYTPAAPGNLVPVCWGK-----GA 57

Db 21 LLLLLLQETGAQDVRVRLPEVRGRLGTVELPCHLLPPTTERVSQVTWQRLLDGTVAAF 80

58 CPVFECGNVLRTERDVNWTSKYWLNGDFRKGDVSLTIENVTLADSGIYCCRI-QIPG 116

Db 81 HPSFGVDFPNSQSKDRLSFVRAPETNADLR--DATAFRGLRVEDEGNVTCFATFPN 138

QY 117 IMDEKENLKLVIKP-----AKVTPAPTLQRFDTAFPRMLTTRGHGPAETQTLGLSPD 170

Db 139 GTRRGVTWLRVIAQEPENHAEAEQVTTIGPQ-----SVAVARCVSTGGRPPA----- 183

171 INLTOSTLANELRDSRLANDLRDSGATIRIGIYIGAGICAGLALALIFGALIFKWSHS 230

```
Db      : | | : | : | :  
184 -RTWISSIGGEAKD-----EPCIOAG----- 206
```

OV 231 KEKTONLSLISLANIPPSGIANAVAEGIRSEENITYTTEENVYEVEEP 277

Db 207 -----TVTTISRYSLVPVGRADGVKVTCPVHESEF-----EFP 239

## RESULT 4

poliovirus receptor mpvr - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
 C/Accession: A53437  
 R/Aoki, J., Koike, S., Ise, I., Sato-Yoshida, Y., Nomoto, A.  
 J. Biol. Chem. 269, 8431-8438, 1994  
 A/Title: Amino acid residues on human poliovirus receptor involved in interaction with F  
 A/Reference number: A53437, MUID:94179228; PMID:8132569  
 A/Accession: A53437  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-530 <AOK>  
 A/Cross-references: GB:J26107; MID:9475017; PIDN:BA05103.1; PID:9825507  
 A/Experimental source: C57/BL6, brain  
 A/Note: Sequence extracted from NCBI backbone (NCBI:146664, NCBI:146667)  
 C/Superfamily: poliovirus receptor; immunoglobulin homology  
 F/47-133/Domain: immunoglobulin homology <IMM>

Query match	6.9%;	Score 109;	DB 2;	Length 530;
-------------	-------	------------	-------	-------------

Matches: 63; Conservative: 33; Mismatches: 103; Indels: 88; Gaps: 10;

11 LLLLLLTSSEV-----EYRAEVQONAYLPCTPTPAAPGNLVPVCMGK-----GA 57

21 LLLLLQETGAQDVRVRLPEVRGRLGGTVELPCHLLPPTTERVSQVTWQRLDGTVAAAF 80

58 CPVFECGNVLTDERDVNYWTSRYWLNQDFRKGDVSLTIENVTLADSGIYCCRI-QIPG 116

81 HPSFGVDFPNSQFSKDRLSFVARPETNADLR--DATAFRLGRVEDEGNYTCEFATFPN 138

117 IMDEKENLKLVIKP-----AKVTPAFTLORDFTAAFPRLTTRGHGAETOTLGLPD 170

139 GTRBGVTLRLVIAOPENHAEAOEYTCIPQ-----SVAVARCVSTGRPPA----- 183

171 INLTOSTLANEIPDSRIANDLRDSCATIRIGIYTGAGTCAGLALALIEGALIFKWSHS 230

```

      :| | :| :| :|
184 -BTWTSIGGEAKTMO-----EPGTOAG----- 206

```

231 KERTONT.SI.TSI.ANT.PBSGI.ANVAEGT.PSEENTV.TIENNVEVEEP 277

[illegible]

## RESULT 5

**S24066**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2003, 05:17:26 ; Search time 24.8247 Seconds  
(without alignments)  
1166.045 Million cell updates/sec

Title: US-10-004-633-24

Perfect score: 1587

Sequence: 1 MSHLPFDCCVLLLLLLLTTR.....CVSSRQOPQPLGCRFAMP 301

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256	16.1	451	2	S71754 cellular hepatitis
2	120	7.6	365	2	JC7780 coxsackie- and ade
3	109	6.9	467	1	HLMSB2 poliovirus recepto
4	109	6.9	530	2	A53437 poliovirus recepto
5	108	6.8	1138	1	S24066 protein-tyrosine k
6	107	6.7	310	1	HNWZRA hemagglutinin prec
7	102.5	6.5	523	2	I50478 neuramin - goldfiss
8	101.5	6.4	478	2	I53960 PR2 alpha - human
9	101.5	6.4	538	2	I68093 PR2 delta - human
10	101	6.4	538	2	UC2457 vascular cell adhe
11	99	6.2	818	1	S44098 brain-derived neur
12	98	6.2	186	2	I61783 sodium channel bet
13	98	6.2	526	2	S70587 butyrophilin precu
14	97.5	6.1	315	1	HNWZVT hemagglutinin prec
15	96	6.0	313	2	T28598 hypothetical prote
16	96	6.0	520	1	S44099 brain-derived neur
17	96	6.0	321	1	T09058 butyrophilin homol
18	95	6.0	503	2	JC5287 SHP substrate-1 pr
19	94.5	6.0	662	2	T16525 hypothetical prote
20	94.5	6.0	773	1	QRRBG secretory componen
21	94	5.9	215	2	A57843 sodium channel bet
22	94	5.9	246	2	A47712 myelin/oligodendro
23	94	5.9	247	2	A55717 myelin/oligodendro
24	93	5.9	313	2	JQ1862 31r protein - vari
25	93	5.9	313	2	H36854 hemagglutinin - va
26	93	5.9	318	2	F72171 K9r protein - vari
27	93	5.9	372	2	C39371 Ig V-region-like B
28	92.5	5.8	309	2	B-lymphocyte activ
29	92	5.8	1091	2	S01998 contactin precurs

30	92	5.8	1134	1	JN0711 protein-tyrosine k
31	92	5.8	5175	2	T20992 hypothetical prote
32	92	5.8	5198	2	T43290 hemisceritin precu
33	91.5	5.8	307	2	S55596 hypothetical prote
34	91.5	5.8	328	2	AH3342 g1px protein (lipo
35	91.5	5.8	459	2	A46254 CD4 precursor - ra
36	91	5.7	344	2	A27681 nonspecific cross-
37	91	5.7	1136	2	A27681 hypothetical prote
38	90.5	5.7	349	2	A34815 carcinomembryonic a
39	90.5	5.7	458	2	S23969 cell-adhesion mole
40	90.5	5.7	458	2	S68177 C-Cam2a protein is
41	90.5	5.7	519	2	A44783 ecto-ATPase precu
42	89.5	5.6	314	1	JQ1793 hemagglutinin prec
43	89.5	5.6	314	1	HNWZVT hemagglutinin prec
44	89	5.6	1327	2	T09402 immunoglobulin-lik
45	88.5	5.6	136	2	I46635 rearranged T-cell

## ALIGNMENTS

## RESULT 1

S71754 cellular hepatitis A receptor HAVcr-1 precursor - green monkey  
N/Alternate names: surface glycoprotein  
C/Species: Cercopithecus aethiops (green monkey, grivet)  
C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 05-Jun-1998  
C/Accession: S71754  
R/Kaplan, G.; Totsuka, A.; Thompson, P.; Akatsuka, T.; Morisugu, Y.; Feinstein, S.M.  
EMBO J. 15, 4282-4296, 1996  
A/Title: Identification of a surface glycoprotein on African green monkey kidney cells  
A/Reference number: S71754; MUID:97015129; PMID:8861957  
A/Accession: S71754  
A/Molecule type: mRNA  
A/Residues: 1-451 <KAP>  
A/Cross-references: EMBL:X98252; NID:G1526573; PID:e247449; PID:G1526574  
A/Experimental source: Kidney  
C/Keywords: glycoprotein  
F.1-17/Domain: signal sequence #status predicted <SIG>  
F.18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match	16.1%; Score 256; DB 2; Length 451;
Best Local Similarity	22.4%; Pred. No. 4e-14;
Matches	101; Conservative 38; Mismatches 123; Indels 188; Gaps 12;
QY	4 HLPFCVLLLLLLLSSEVEYRAVQNAVLPFFTPAAGNVLPVCKGKACPVREC 63
DB	2 HLQVVLSTLHLADSVADVNDGAGLSTLPCRNV---GALTSCMKNRGTCVSFSC 57
QY	64 GNVVLRTERDVNY-WTSRYWLGDFPKGVSITFENVTLADSGIYCCRIQIPGIMNDEK 122
DB	58 PDGIWVINGTHTVTRKETRYKLGNLSRRDVSLLIANTAVSDSGIYCCRVHSGFNDMK 117
QY	123 FNLKLVKPAKV----- 134
DB	118 ITISLKIGPRVTVPIVATVSTVPTTLLPTTLLPTTLLPTTLLPTTLLPTMT 177
QY	135 ----- 134
DB	178 TLPPTTVPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLL 237
QY	135 -----TPAPTQORDTAFPRMLT-----TRGH 157
DB	238 PTTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLLPTTLL 297
QY	158 GPATQTGSLPDNLQISTLANELR-----DSRLANDLRDSCATI----- 199
DB	298 EPVATSP--SSPGQAEHNPVLLGATRTQPTSSSLYKSTTDGSDTVSSDGLMNNQTO 355
QY	200 -----RIGIYIGAGICAG-LALALIFGALIFKWSKSEKIQNISLSIANLP 246
DB	356 LSPHSQPMVNTTEGIY--AGVCISVILAVLAVGLVIAKKYFFKKE-IQQLS-VSFSHQ 411



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:56:41 ; Search time 3372.32 Seconds  
(without alignments)  
10954.308 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903

Sequence: 1 atgttcacatctccctt.....gtgtcgtcttgcattgcca 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sfs:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pac:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sfs:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rtd:\*

36: em\_hcg\_mam:\*

37: em\_hcg\_vrt:\*

38: em\_sy:\*

39: em\_hcgo\_hum:\*

40: em\_hcgo\_mus:\*

41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	903	100.0	903 6	AR143581
2	903	100.0	903 6	AR168950
3	903	100.0	903 6	AR232698
4	903	100.0	906 9	AY069944
5	903	100.0	1116 9	AF450242
6	903	100.0	1772 6	BD157433
7	903	100.0	1772 6	AK027334
8	903	100.0	2236 6	AR143568
9	903	100.0	2236 6	AR168937
10	903	100.0	2236 6	AR232685
11	901.4	99.8	1116 9	AF450243
12	901.4	99.8	2320 9	AF251707
13	712	78.8	769 6	BD146047
14	456	50.5	2710 6	AR143561
15	456	50.5	2710 6	AR168930
16	456	50.5	2710 6	AR232678
17	456	50.5	2710 6	189415
18	456	50.5	2725 10	AF450241
19	455	50.4	529 6	BD060055
20	442.6	49.0	843 6	AR143567
21	442.6	49.0	843 6	AR168936
22	442.6	49.0	843 6	AR232684
23	440.8	48.8	862 10	AF39831
24	396.6	43.9	1012 9	BC020843
25	374.2	41.4	438 6	BD076086
26	340.6	37.7	54375 5	AC011377
27	211.4	23.4	177087 2	AC119505
28	211.4	23.4	233370 2	AC098539
29	187.6	20.8	223030 10	AL669948
30	148	16.4	228827 2	AC135694
31	140.6	15.6	223030 10	AL669948
32	129.4	14.3	233370 2	AC098539
33	109.2	12.1	177087 2	AC119505
34	74	8.2	1507 9	BC013325
35	72.6	8.0	1440 9	AF043724
36	72.6	8.0	1795 6	BD062750
37	67.6	7.5	39980 9	AC011550
38	67.6	7.5	106169 9	AC073534
39	66.4	7.4	918 10	AF399829
40	66.4	7.4	2331 10	BC032879
41	66	7.3	140330 9	AC008491
42	65.4	7.2	143661 9	AC008491
43	65.4	7.2	849 10	AF399830
44	65.4	7.2	1893 10	BC03400
45	65.4	7.2	197196 2	AC112585

ALIGNMENTS

RESULT 1

LOCUS AR143581 903 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 37 from patent US 6204371.

ACCESSION AR143581

VERSION AR143581.1 GI:15104867

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 903)

AUTHORS Levinson,D.Adam.

TITLE Compositions and methods for the treatment and diagnosis of immune disorders

JOURNAL Patent: US 6204371-A 37 20-MAR-2001;

FEATURES		Location/Qualifiers					
	SOURCE	1..903 /organism="unknown"					
	BASE COUNT	242 a 225 c 221 g 215 t					
	ORIGIN						
<b>Query Match 100.0%; Score 903; DB 6; Length 903;</b>							
<b>Best Local Similarity 100.0%; Pred. No. 1.1e-277;</b>							
<b>Matches</b>	<b>903; Conservative</b>	<b>0;</b>	<b>Mismatches</b>	<b>0;</b>	<b>Indels</b>	<b>0;</b>	<b>Gaps</b>
OY	1 ATGTTTACACTCTCCCTTGAGCTGTGCCTGCTGCTCTCTACTACTACAAAG	60					
Db	1 ATGTTTCAACNTCCCTTTGACGTGTGCCTGCTGCTGCTCTACTACTTAACAAG	60					
OY	61 TCCTCAGAAGTGAATPACAGACGGAGGTGGTCAGAAATGCTATCTGCCCTCTTAC	120					
Db	61 TCCTCAGAAGTGAATPACAGACGGAGGTGGTCAGAAATGCTATCTGCCCTCTTAC	120					
OY	121 ACCCAAGCCCCCGAAGAACCTCGTCCCTGTCTGTGGGGCAAAAGACCCTGTCTG	180					
Db	121 ACCCAAGCCCCCGAAGAACCTCGTCCCTGTCTGTGGGGCAAAAGACCCTGTCTG	180					
OY	181 TTYGAATGNGCAAGTGTGTGCTCAGAGACTGATGAAGGAGTGAATTATGACATCC	240					
Db	181 TTYGAATGNGCAAGTGTGTGCTCAGAGACTGATGAAGGAGTGAATTATGACATCC	240					
OY	241 AGATACTGGCTAATATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACATAGAAATGTG	300					
Db	241 AGATACTGGCTAATATGGGGATTTCCGCAAAAGAGATGTGTCCCTGACATAGAAATGTG	300					
OY	301 ACTCTAGCAGACAGTGGGATCTTACGTCTGCCGATCCAAATCCCAAGGCTAATGAATGAT	360					
Db	301 ACTCTAGCAGACAGTGGGATCTTACGTCTGCCGATCCAAATCCCAAGGCTAATGAATGAT	360					
OY	361 GAAAAATTTAACCTGAAGTGTGATCAAACCGACCAAGGTCAACCCCTGCACCGACTCTG	420					
Db	361 GAAAAATTTAACCTGAAGTGTGATCAAACCGACCAAGGTCAACCCCTGCACCGACTCTG	420					
OY	421 CAGAGAGACTTCACTGCAGACCCTTTCMAAGATGCTTACCAACGAGGGACATGGCCACGA	480					
Db	421 CAGAGAGACTTCACTGCAGACCCTTTCMAAGATGCTTACCAACGAGGGACATGGCCACGA	480					
OY	481 GAGACACAGACACTGGGGAGCCTCCCTGATATAAATCTAACCAAAATATCCAATTGGCC	540					
Db	481 GAGACACAGACACTGGGGAGCCTCCCTGATATAAATCTAACCAAAATATCCAATTGGCC	540					
OY	541 AATGAGTTACGGGACCTCTGATTTGGCAATGAATTACGGGACCTGGAGACCAATACGA	600					
Db	541 AATGAGTTACGGGACCTCTGATTTGGCAATGAATTACGGGACCTGGAGACCAATACGA	600					
OY	601 ATAAGCATCTACATCGAGACAGAGACTGTGTGGCTGGCTCTGGCTTTATCTTCCGCG	660					
Db	601 ATAAGCATCTACATCGAGACAGAGACTGTGTGGCTGGCTCTGGCTTTATCTTCCGCG	660					
OY	661 GCCTTAATTTCAATGTGATTTCTCATAGCAAAAGAGATACAGATTTAAAGCTCATC	720					
Db	661 GCCTTAATTTCAATGTGATTTCTCATAGCAAAAGAGATACAGATTTAAAGCTCATC	720					
OY	721 TCTTTGGCCACTCCCTCTCCCTCAGGATTTGGCAATGACGTATGACAGAGGAATTTGGCTCA	780					
Db	721 TCTTTGGCCACTCCCTCTCCCTCAGGATTTGGCAATGACGTATGACAGAGGAATTTGGCTCA	780					
OY	781 GAAGAAAACATCTATACATTGAAGAGACGTATATGAAGTGAAGAGCCCAATAGTAT	840					
Db	781 GAAGAAAACATCTATACATTGAAGAGACGTATATGAAGTGAAGAGCCCAATAGTAT	840					
OY	841 TAATGCTATGACAGACAGGACCAACCTTCAACAACCTTTGGGTTGTGCTTTGCAATG	900					
Db	841 TAATGCTATGACAGACAGGACCAACCTTCAACAACCTTTGGGTTGTGCTTTGCAATG	900					
OY	901 CCA 903						
Db	901 CCA 903						

[illegible]

Dd		661	GCTTTAATTTTTCAAAATGGTATTCTCATAGCANAAGAATACAGATTTTAAGCCATAC	720
Oy		721	TCTTTGGCCAACCTCCCTTCCTCAGAGTTGGCAAATGCAGTAGCAGAGGAATTGCTCA	780
Dd		721	TCTTTGGGCCAACCTCCCTTCCTCAGAGTTGGCAAATGCAGTAGCAGAGGAATTGCTCA	780
Oy		761	GAAGAAAACATCATATACCATTTGAAGAAGACGTATATGAAGTGAGAGGCCCAATGAGTAT	840
Dd		761	GAAGAAAACATCATATACCATTTGAAGAAGACGTATATGAAGTGAGAGGCCCAATGAGTAT	840
Oy		841	TATTGCTATGTGTCAGCAGCAGCAGCAACCCTACAACCTTTGGTTGTGCTTTGCAATG	900
Dd		841	TATTGCTATGTGTCAGCAGCAGCAGCAACCCTACAACCTTTGGTTGTGCTTTGCAATG	900
Oy		901	CCA 903	
Dd		901	CCA 903	
RESULT 3				
LOCUS		AR332698	903 bp	DNA linear PAT 20-DEC-2002
DEFINITION		Sequence 37 from patent US 6455685.		
ACCESSION		AR332698		
VERSION		AR332698.1	GI:27274975	
KEYWORDS				
SOURCE		Unknown.		
ORGANISM		Unknown.		
REFERENCE		1 (bases 1 to 903)		
AUTHORS		Levineon,D.A.		
TITLE		Compositions and methods for the treatment and diagnosis of immune disorders		
JOURNAL		Patent: US 6455685-A 37 24-SEP-2002;		
FEATURES		Location/Qualifiers		
source		1..903		
BASE COUNT		/organism="unknown"		
ORIGIN		242 a 225 c 221 g 215 t		
Query Match		100.0%; Score 903; DB 6; Length 903;		
Best Local Similarity		100.0%; Pred.No.1.1e-277;		
Matches		903; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Oy		1	ATGTTTTCACATCTTCCCTTGTACTGTGTCTGCTGCTGCTGCTACTACTTACAAG	60
Dd		1	ATGTTTTCACATCTTCCCTTGTACTGTGTCTGCTGCTGCTGCTGCTACTACTTACAAG	60
Oy		61	TCCTCAGAAATGGAATPACAGACCGGAGGTGGTCAAGATCCTATATCGCTGCTTTAC	120
Dd		61	TCCTCAGAAATGGAATPACAGACCGGAGGTGGTCAAGATCCTATCTGCTGCTTTAC	120
Oy		121	ACCCGAGCGGCCCGAGGGAACTCGTCCCGCTCTGCGGCAAGAGACCTGTCTGTG	180
Dd		121	ACCCGAGCGGCCCGAGGGAACTCGTCCCGCTCTGCGGCAAGAGACCTGTCTGTG	180
Oy		181	TTTGAATGTGCGACGTTGTGCTCAGACTGATGAAGGATGTGAATTTATTTGACATCC	240
Dd		181	TTTGAATGTGCGACGTTGTGCTCAGACTGATGAAGGATGTGAATTTATTTGACATCC	240
Oy		241	AGATACTGGCTAAATGCGGATTTCCGCAAGAGATGTGTCCCTGACCAATAGAAATGTG	300
Dd		241	AGATACTGGCTAAATGCGGATTTCCGCAAGAGATGTGTCCCTGACCAATAGAAATGTG	300
Oy		301	ACTTAGCAGACAGTGGGATCTACTGCTGCGCGGATCCAAATGCCAGGCAATATGAATAT	360
Dd		301	ACTTAGCAGACAGTGGGATCTACTGCTGCGCGGATCCAAATGCCAGGCAATATGAATAT	360
Oy		361	GAATAATTTAACTGTAAGTTGTGTATCAAAACGACCAAGTCAACCCCTGACCGACTGTG	420
Dd		361	GAATAATTTAACTGTAAGTTGTGTATCAAAACGACCAAGTCAACCCCTGACCGACTGTG	420
Oy		421	CAGAGAGACTTACTGACAGCCCTTTCCAAAGATGCTTACCAACGAGGACATGGCCACAGA	480

Db	421	CAGAGAGCTTCACTGCAGCGCTTTCAGAGATGCTTACACACAGGGGACATGGCCACGA	480
Qy	481	GAGACACAGACACTGGGGAGCCCTCCGTATTAATCTAACACAAATATCCATTGGCC	540
Db	481	GAGACACAGACACTGGGGAGCCCTCCGTATTAATCTAACACAAATATCCATTGGCC	540
Qy	541	AATGATTAACGGGACTCTAGATTGGCCCATGACTTTACGGGACTCTGGAGCAACATCAGA	600
Db	541	AATGATTAACGGGACTCTAGATTGGCCCATGACTTTACGGGACTCTGGAGCAACATCAGA	600
Qy	601	ATTAGGCATCTACATCCGAGCGAGGATATGTGCTGTGGGCTGGCTTATCTTGGCC	660
Db	601	ATTAGGCATCTACATCCGAGCGAGGATATGTGCTGTGGGCTGGCTTATCTTGGCC	660
Qy	661	GCTTTAATTTTCAATGGTATTTCTATAGCAAAAGAAATACAGAAATTTAAGCCTATC	720
Db	661	GCTTTAATTTTCAATGGTATTTCTATAGCAAAAGAAATACAGAAATTTAAGCCTATC	720
Qy	721	TCCTTTGGCCAACTCCCTCCCTCAGATTGGCAATGACATGACAGAGGGAATTCGCTCA	780
Db	721	TCCTTTGGCCAACTCCCTCCCTCAGATTGGCAATGACATGACAGAGGGAATTCGCTCA	780
Qy	781	GAAAGAAAACATCTATTACCATTTGAAGAGAAAGTAAATGAATGACGAGGCCCAATGAGTAT	840
Db	781	GAAAGAAAACATCTATTACCATTTGAAGAGAAAGTAAATGAATGACGAGGCCCAATGAGTAT	840
Qy	841	TATTGCTATGTACAGACAGGACAGCAACCTCACAACCTTTGGGTTGTGCTTTGCATG	900
Db	841	TATTGCTATGTACAGACAGGACAGCAACCTCACAACCTTTGGGTTGTGCTTTGCATG	900
Qy	901	CCA 903	
Db	901	CCA 903	
RESULT 4			
LOCUS	AY069944	906 bp	mRNA linear PRI 01-JUL-2002
DEFINITION	Homo sapiens putative kidney injury molecule-3 mRNA, complete cds.		
ACCESSION	AY069944		
VERSION	AY069944.1	GI:21655105	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 906) Kuehn,E.W., Ichimura,T. and Bonventre,J.V. A homolog to human kidney injury molecule-1 is expressed in hepatoma cells		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 906) Kuehn,E.W., Ichimura,T. and Bonventre,J.V. Direct Submission		
AUTHORS	Submitted (11-DEC-2001) Renal Unit, Massachusetts General Hospital, 149, 13th Street, Charlestown, MA 02129, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1..906		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="5"		
	/cell_line="HepG2"		
	/tissue_type="hepatocellular carcinoma"		
	1..906		
	/note="hKIM-3; similar to Homo sapiens kidney injury molecule-1 (hKIM-1)"		
	/codon_start=1		
	/product="putative kidney injury molecule-3"		
	/protein_id="AA155401.1"		
	/db_xref="GI:21655106"		
	/translation="MFSHLPFDVCLLLLLLIRSSSEVRAEVGQNAVYPCFYTPAA		

BASE COUNT 243 a 225 c 222 g 216 t

ORIGIN

PNNVLPVCMWKGACAFVECCGAVVLRTEREDVNVYTSRMYLNDGDERKGDVSLTTEENTL  
 ADGTCVCCRIQIPGIMNDEKFNKLTKVAPKATPAPTORFPAAPMLTTRHGPAA  
 ETOHGLSPDLNLTQISTLANELDRSLANDLRGSGATIRIGITGAGTCAGLALAIL  
 FGALEFKWYSHSKERKIONLTLSISLANLPGLSANAVAAGTISEENITYTIEBNVYVEEBE  
 PNEYCYVUSSRQOSPDLGCRFMPP"

Query Match	100.0%	Score 903	DB 97	Length 906
Best Local Similarity	100.0%	Pred. No. 1.1e-277		
Matches 903	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGTTTTCACATCTTCCCTTTGACTGTGTCTGTGCTGCTGCTACTACTTACAAAG	60	
Db	1	ATGTTTTCACATCTTCCCTTTGACTGTGTCTGTGCTGCTGCTACTACTTACAAAG	60	
QY	61	TCCTCAGAACTGGAATACAGACCGGAGGTGGGTGAGAAATGCTTATCTGCTGCTTAC	120	
Db	61	TCCTCAGAACTGGAATACAGACCGGAGGTGGGTGAGAAATGCTTATCTGCTGCTTAC	120	
QY	121	ACCCGAGCGCGCCCGAGGGAACTCGTGGCCGCTGCTGCTGAGGCAAAAGAGCCTGTCTGTG	180	
Db	121	ACCCGAGCGCGCCCGAGGGAACTCGTGGCCGCTGCTGCTGAGGCAAAAGAGCCTGTCTGTG	180	
QY	181	TTTGAATGTGGCAAGTGTGTCTCAGACTGATGAAGAGATGTGAATTTATTTGACATCC	240	
Db	181	TTTGAATGTGGCAAGTGTGTCTCAGACTGATGAAGAGATGTGAATTTATTTGACATCC	240	
QY	241	AGATTTCTGGCTTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGTACCATAGAGATGTG	300	
Db	241	AGATTTCTGGCTTAAATGGGGATTTCCGCAAAAGAGATGTGTCCCTGTACCATAGAGATGTG	300	
QY	301	ACTCTAGCAGACAGTGGGATCTACTGCTGCGCCGAGTCCAAATCCGAGCATATGATGAT	360	
Db	301	ACTCTAGCAGACAGTGGGATCTACTGCTGCGCCGAGTCCAAATCCGAGCATATGATGAT	360	
QY	361	GAATAATTTAACTCTGAAGTTGGTATCAAAACGAGCAAGATGACCCCTGCACGACTCTG	420	
Db	361	GAATAATTTAACTCTGAAGTTGGTATCAAAACGAGCAAGATGACCCCTGCACGACTCTG	420	
QY	421	CAGAGAGCTTCACTGCAAGCCTTTCCAGAGATGCTTACCACCAAGGGGCAATGGCCACGA	480	
Db	421	CAGAGAGCTTCACTGCAAGCCTTTCCAGAGATGCTTACCACCAAGGGGCAATGGCCACGA	480	
QY	481	GAGACACAGACACTGGGGAGCCTCCCTGATATTAATTTAACAACAATATCCACATTTGGCC	540	
Db	481	GAGACACAGACACTGGGGAGCCTCCCTGATATTAATTTAACAACAATATCCACATTTGGCC	540	
QY	541	AATGAGTTACGGGACTCTAGATTTGGCCAAATGACTTACCGGACTCTGAGCAACCATCAGA	600	
Db	541	AATGAGTTACGGGACTCTAGATTTGGCCAAATGACTTACCGGACTCTGAGCAACCATCAGA	600	
QY	601	ATAGGCATCTTCAATGGGACAGGATCTGTGTCTGGGCTGGGCTCTGGCTCTTATCTTGGGC	660	
Db	601	ATAGGCATCTTCAATGGGACAGGATCTGTGTCTGGGCTGGGCTCTGGCTCTTATCTTGGGC	660	
QY	661	GCTTAAATTTTCAATGTGTAATTTCTATAGCAAAAGAGATACAGAAATTTAAGCCTCATC	720	
Db	661	GCTTAAATTTTCAATGTGTAATTTCTATAGCAAAAGAGATACAGAAATTTAAGCCTCATC	720	
QY	721	TCCTTGGCCAACTTCCCTCAGAGATTTGGCAATGCAATGCAAGAGGGAATTCGCTCA	780	
Db	721	TCCTTGGCCAACTTCCCTCAGAGATTTGGCAATGCAATGCAAGAGGGAATTCGCTCA	780	
QY	781	GAAGAAAACATCTATATCCATTGAAGAAAGTAATTTGAATGTGGAGAGCCCAATGAGAT	840	
Db	781	GAAGAAAACATCTATATCCATTGAAGAAAGTAATTTGAATGTGGAGAGCCCAATGAGAT	840	
QY	841	TATTTGCTATGTGACAGACAGGCAAGCCCTCACAACCTTTGGGTGTGCGCTTTGCATG	900	
Db	841	TATTTGCTATGTGACAGACAGGCAAGCCCTCACAACCTTTGGGTGTGCGCTTTGCATG	900	
QY	901	CCA 903		

Db	901	CCA	903
RESULT 5	AF450242	1116 bp	RNA linear PRI 11-FEB-2002
LOCUS	AF450242		
DEFINITION	Homo sapiens clone 1 T cell immunoglobulin mucin-3 (TIM3) mRNA,		
ACCESSION	AF450242		
VERSION	AF450242.1	GI:18182532	
KEYWORDS			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Catarrhini; Hominiidae; Homo. Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1116)		
TITLE	Monney,L., Sabatos,C.A., Gaglile,J.L., Ryu,A., Waldner,H., Chernova,T., Manning,S., Greenfield,E.A., Coyle,A.J., Sobel,R.A., Freeman,G.J. and Kuchroo,V.K. Tni-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease Nature 415 (6871), 536-541 (2002)		
JOURNAL	Nature 415 (6871), 536-541 (2002)		
MEDLINE	21681868		
PUBMED	11823861		
REFERENCE	2 (bases 1 to 1116)		
AUTHORS	Monney,L., Sabatos,C., Gaglile,J.L., Ryu,A., Waldner,H., Chernova,T., Greenfield,E.A., Sobel,R.A., Freeman,G.J. and Kuchroo,V.K. Direct Submission Submitted (23-NOV-2001) Adult Oncology, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115, USA		
TITLE	Location/Qualifiers		
JOURNAL	1. 1116		
FEATURES	/organism="Homo sapiens"		
source	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="1"		
	/cell_type="T cell"		
	1. 1116		
	/gene="TIM3"		
	58. 963		
	/gene="TIM3"		
	/codon_start=1		
	/product="T cell immunoglobulin mucin-3"		
	/protein_id="AA65157.1"		
	/db_xref="GI:18182533"		
	/translation="MFSHLPFDCVLLHLLHLLTRSSSEVEYRAEYQNAVYPCTYTPAA /PGNIVPCWKGACPVCEGNCVVLRTDERVNVNTSYMLNGDPKGDVSLTIENVTH /ADSGVYCRITQIGIMNDKRNKIVKIPAKTPAPLTIORDFPAAFPRLTTGCHGPF /ETQGLFSLPIPNILTOITSLANEKRLDRSLANDLSDSGATTIRIGIYAGICAGLALAIL /EGQLFELSHSKKIONLSLISLANLPFGLANAVABGRISENVIYTBENVYEEB /PNEYCVSSROOPSPQGLGRFAMP"		
	PNEYCVSSROOPSPQGLGRFAMP"		
BASE COUNT	291 a 268 c 270 g 287 t		
ORIGIN			
Query Match	100.0%;	Score: 903;	DB 9; Length 1116;
Best Local Similarity	100.0%;	Pred. No. 1.1e-277;	
Matches 903;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Db	1	ATGTTTTCACATCTTCCCTTGA	CTGTGTCCTGCTGCTGCTACTACTTACAAAG 60
Qy	58	ATGTTTTCACATCTTCCCTTGA	CTGTGTCCTGCTGCTGCTACTACTTACAAAG 117
Db	61	TCTTCAGAAAGTGAATACAGAC	CGGAGTGGTCAGAATGCTATTCGCTGCTTCTAC 120
Qy	118	TCTTCAGAAAGTGAATACAGAC	CGGAGTGGTCAGAATGCTATTCGCTGCTTCTAC 177
Db	121	ACCCGAGCGCGCCAGGGAACCT	GTCGTCGCTGCGGCAAGAGAGCTGTCCTGTG 180
Qy	178	ACCCGAGCGCGCCAGGGAACCT	GTCGTCGCTGCGGCAAGAGAGCTGTCCTGTG 237
Db	181	TTTGAATGTGGCAACGTGTGCT	CAGAGCTAGTGAAGGATGTGAATTATGGACATCC 240

```

Db      |||||
238 TTTGATGTGGCAAGTGTGCTCAGAGCTGTGAAGGAGTGAATTAATGACATCC 297
Qy      AGATACGTGCTAAATGGGGATTTCCGCAAGAGATGTGTCCCTGACCATAGAGAAATG 300
Db      AGATACGTGCTAAATGGGGATTTCCGCAAGAGATGTGTCCCTGACCATAGAGAAATG 357
Qy      ACTCTAGACAGACAGTGGGATCTACTGTCCGCGATCCAAATCCAGAGATATGATAT 360
Db      ACTCTAGACAGACAGTGGGATCTACTGTCCGCGATCCAAATCCAGAGATATGATAT 417
Qy      GAAAAATTTAACTCTGAATGTGCTCAAAACGAGCAAGTCAACCCCTGCAACGACTGTG 420
Db      GAAAAATTTAACTCTGAATGTGCTCAAAACGAGCAAGTCAACCCCTGCAACGACTGTG 477
Qy      CAGAGAGACTCAGTCGAGCCCTTCCAGAGATGCTTACCAAGGGGACATGGCCACGA 480
Db      CAGAGAGACTCAGTCGAGCCCTTCCAGAGATGCTTACCAAGGGGACATGGCCACGA 537
Qy      GAGACACAGACACTGGGAGCCCTCCGTATATAATCTAACACAAATATCCACATTGGCC 540
Db      GAGACACAGACACTGGGAGCCCTCCGTATATAATCTAACACAAATATCCACATTGGCC 597
Qy      AATGATTAACGGGACTAGATTGGCAATGACTTACGGGACTGTGAGCAACATCAGA 600
Db      AATGATTAACGGGACTAGATTGGCAATGACTTACGGGACTGTGAGCAACATCAGA 657
Qy      ATAGGCACTACATCGGAGCAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 660
Db      ATAGGCACTACATCGGAGCAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 717
Qy      GCTTAAATTTTCAATGTGATTTCTCATAGCAAAAGAGATACAGAAATTTAAGCTCTATC 720
Db      GCTTAAATTTTCAATGTGATTTCTCATAGCAAAAGAGATACAGAAATTTAAGCTCTATC 777
Qy      TCTTTGGCCAACTCTCCCTCAGAGATTGGCAATGAGTAGAGAGGAAATTCCTCA 780
Db      TCTTTGGCCAACTCTCCCTCAGAGATTGGCAATGAGTAGAGAGGAAATTCCTCA 837
Qy      GAAAGAAACATCTAATCAATGAAAGAAACGATATGAAATGAGAGAGCCCAATGATAT 840
Db      GAAAGAAACATCTAATCAATGAAAGAAACGATATGAAATGAGAGAGCCCAATGATAT 897
Qy      TATTCCTATGTCTAGCAGAGCAGCAACCTTCAACCTTTGGGTTGTCTGCTTTCATG 900
Db      TATTCCTATGTCTAGCAGAGCAGCAACCTTCAACCTTTGGGTTGTCTGCTTTCATG 957
Qy      901 CCA 903
Db      958 CCA 960

RESULT 6
BD157433      1772 bp      DNA      linear      PAT 17-JAN-2003
LOCUS      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD157433
VERSION      BD157433.1 GI:27863191
KEYWORDS      JP 2002191363-A/12276.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1772)
AUTHORS      Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
TITLE      Ishii,S., Sugiyama,T., Wakematsu,A., Nagai,K. and Otsuki,T.
JOURNAL      Primer for synthesizing full-length cDNA and use thereof
PATENT      Patent: JP 2002191363-A 12276 09-JUL-2002;
COMMENT      HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/12276
PD      09-JUL-2002
PF      28-JUL-2000 JP 2000280990

```

```

PI      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI      SAITO,
PI      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI      KEIICHI NAGAI, TETSUJI OTSUKI
PC      C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC      10,
PC      C12P21/02, C1201/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC      Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source      Location/Qualifiers
FT          CDS      Location/Qualifiers
              1..1772
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT      473 a      396 c      419 g      484 t
ORIGIN

```

```

Query Match      100.0%; Score 903; DB 6; Length 1772;
Best Local Similarity 100.0%; Pred. No. 1,2e-277;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGTTTACATCTTCCCTTGACTGTCTCTGCTGCTGCTGCTACTACTTACAAG 60
Db      44 ATGTTTACATCTTCCCTTGACTGTCTCTGCTGCTGCTGCTACTACTTACAAG 103
Qy      61 TCTTCAGAGTGAATACAGAGCGGAGTCCGTAGAAATGCTTATTCCTCTTCTAC 120
Db      104 TCTTCAGAGTGAATACAGAGCGGAGTCCGTAGAAATGCTTATTCCTCTTCTAC 163
Qy      121 ACCCGACCGCCCGCAGGGAACCTGTGCGCTGCTGGGGCAAGAGCGCTGCTGTG 180
Db      164 ACCCGACCGCCCGCAGGGAACCTGTGCGCTGCTGGGGCAAGAGCGCTGCTGTG 223
Qy      181 TTTGAATGTGCAACGTGTGCTCAGAGCTGATGAAAGGATGTAATTTATGACATCC 240
Db      224 TTTGAATGTGCAACGTGTGCTCAGAGCTGATGAAAGGATGTAATTTATGACATCC 283
Qy      241 AGATACGTGCTAAATGGGGAATTTCCGCAAGAGATGTCTCTGACCATAGAAATG 300
Db      284 AGATACGTGCTAAATGGGGAATTTCCGCAAGAGATGTCTCTGACCATAGAAATG 343
Qy      301 ACTCTAGACAGACAGTGGGATCTACTGCTCCGAGATCCAAATCCAGGCAATGATAT 360
Db      344 ACTCTAGACAGACAGTGGGATCTACTGCTCCGAGATCCAAATCCAGGCAATGATAT 403
Qy      361 GAAAAATTTAACTCTGAATGTGCTCATCAACCAAGGCTACCCCTGACCGACTCTG 420
Db      404 GAAAAATTTAACTCTGAATGTGCTCATCAACCAAGGCTACCCCTGACCGACTCTG 463
Qy      421 CAGAGAGCTTCACTGCAAGCTTTCCAGAGATGCTTTCACACGAGGAGCATGGCCACGA 480
Db      464 CAGAGAGCTTCACTGCAAGCTTTCCAGAGATGCTTTCACACGAGGAGCATGGCCACGA 523
Qy      481 GAGACACAGACACTGGGAGCTCCGTATTAATCTTAACAACAATATCCAAATGGCC 540
Db      524 GAGACACAGACACTGGGAGCTCCGTATTAATCTTAACAACAATATCCAAATGGCC 583
Qy      541 AATGATTAACGGGACTCTAGATTGGCAATGACTTACGGGACTGTGAGCAACATCAGA 600
Db      584 AATGATTAACGGGACTCTAGATTGGCAATGACTTACGGGACTGTGAGCAACATCAGA 643
Qy      601 ATAGGCACTACATCGGAGCAGGAGATCTGTGCTGGGCTGTGCTTATCTTGGCC 660
Db      644 ATAGGCACTACATCGGAGCAGGAGATCTGTGCTGGGCTGTGCTTATCTTGGCC 703
Qy      661 GCTTAAATTTTCAATGTGATTTCTCATAGCAAAAGAGATACGAATTTAAGCTCATC 720
Db      704 GCTTAAATTTTCAATGTGATTTCTCATAGCAAAAGAGATACGAATTTAAGCTCATC 763
Qy      721 TCTTTGGCCAACTCTCCCTCAGAGATTGGCAATGACAGAGGAAATTCGCTCA 780

```

Db 764 TCTTTGGCCAACTCCCTCCCTCAGGATTGGCAATGCGATGACAGAGGGAATTCCTCTCA 823

QY 781 GAAGAAAACATCTATACCATTTGAAGACGTATATGAAGTGAGAGGCCCAATGAT 840

Db 824 GAAGAAAACATCTATACCATTTGAAGACGTATATGAAGTGAGAGGCCCAATGAT 883

QY 841 TATGTCTATGTGACGACGAGGCAACCTCTCAACCTTTGGTTGTGCTTTGCAAG 900

Db 884 TATGTCTATGTGACGACGAGGCAACCTCTCAACCTTTGGTTGTGCTTTGCAAG 943

QY 901 CCA 903

Db 944 CCA 946

RESULT 7

LOCUS AK027334 1772 bp mRNA linear PRI 01-AUG-2002

DEFINITION Homo sapiens cDNA FLJ14428 fis, clone HEMBA1006293.

ACCESSION AK027334

VERSION AK027334.1 GI:14041941

KEYWORDS oligo capping, file (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanabe, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuno, Y. and Oshima, A.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1772)

AUTHORS Isogai, T. and Otsuki, T.

TITLE Direct Submission

JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 282-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'-3'-end one pass selection and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

source

1..1772

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="HEMBA1006293"

/issue\_type="whole embryo, mainly head"

/clone\_lib="HEMBA1"

/dev\_stage="embryo, 10 weeks"

/note="Cloning vector: pMT185TL3"

44..949

/note="unnamed protein product"

/codon\_start=1

/protein\_id="BAB55044.1"

/db\_xref="GI:14041942"

/translation="MESHLPDFCVLLLLLLLRSEVEYRAVGVNATLPCTTPAA

PGMLVPVCKGKACVPFCGNVLTDERVNTWISRYLNGDFRKGDSLITENVTL

ADSGIYCCRIQIPGINNDEKFNKLVIKPAKVPAPLQDFPAAPRMLTTRGHGPA

ETQTLGSLPDINTQISTLANELDRSLANDLSDGATIRIGIYAGICAGIALALI

FGALIPKMYSHSKKQISLISIANIPSGLANAVAGGIRSENIYITEENVYEEB

PNEYCYVSSRQPSQDLCGRFAMP"

BASE COUNT 473 a 396 c 419 g 484 t

ORIGIN

Query Match 100.0%; Score 903; DB 9; Length 1772;

Best local similarity 100.0%; Pred. No. 1.2e-277;

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTCACTCTTCCCTTGTACTGTGTCTGTCTGTCTGTCTACTTACACAG 60

Db 44 AGTTTTCACATCTTCCCTTGTACTGTGTCTGTCTGTCTGTCTACTTACAG 103

QY 61 TCCCTAGAAAGTGAATPACAGAGCGGAGGTCGAGAAATGCTATCTGCGCTTAC 120

Db 104 TCCCTAGAAAGTGAATPACAGAGCGGAGGTCGAGAAATGCTATCTGCGCTTAC 163

QY 121 ACCCGAGCGCCCGCCAGGGAACCTCGTGCCTGTCTGTCTGTCTGTCTGTG 180

Db 164 ACCCGAGCGCCCGCCAGGGAACCTCGTGCCTGTCTGTCTGTCTGTCTGTG 223

QY 181 TTTGAATGTGGCAAGTGTGTCTCAGACTGATGAAGGATGTGAATTTATGACATC 240

Db 224 TTTGAATGTGGCAAGTGTGTCTCAGACTGATGAAGGATGTGAATTTATGACATC 283

QY 241 AGATACGTGGCTAAATGGGGATTTCCGCAAGGAGATGTGTCTGACCATAGAAATGTC 300

Db 284 AGATACGTGGCTAAATGGGGATTTCCGCAAGGAGATGTGTCTGACCATAGAAATGTC 343

QY 301 ACTTAGCAGACAGTGGGATCTTACTGTCTGCGCATCCAAATCCAGGCAATGAATGAT 360

Db 344 ACTTAGCAGACAGTGGGATCTTACTGTCTGCGCATCCAAATCCAGGCAATGAATGAT 403

QY 361 GAAAAATTTAACTGAAATGTGTCTCAAAACGACGCAAGTCAACCTGTGACCTGTG 420

Db 404 GAAAAATTTAACTGAAATGTGTCTCAAAACGACGCAAGTCAACCTGTGACCTGTG 463

QY 421 CAGAGAGACTTCACTGACGAGCTTTCCAGAGATGCTTACCAAGGAGGATCATGAGCC 480

Db 464 CAGAGAGACTTCACTGACGAGCTTTCCAGAGATGCTTACCAAGGAGGATCATGAGCC 523

QY 481 GAGACACAGACACTGGGAGCCTCCCTGATATATTAACAATAATATCAATTTGCC 540

Db 524 GAGACACAGACACTGGGAGCCTCCCTGATATATTAACAATAATATCAATTTGCC 583

QY 541 AATGAGTTACGGGACTCTGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACATCA 600

Db 584 AATGAGTTACGGGACTCTGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACATCA 643

QY 601 ATAGGCATCTACATCGAGACAGGATCTGTGCGGCTGCTGTGCTTATCTTCGCG 660

Db 644 ATAGGCATCTACATCGAGACAGGATCTGTGCGGCTGCTGTGCTTATCTTCGCG 703

QY 661 GCTTTAATTTTCAAAATGTATTTCTCATAGCAAGAGATACAGAAATTTAAGCTCATC 720

Db 704 GCTTTAATTTTCAAAATGTATTTCTCATAGCAAGAGATACAGAAATTTAAGCTCATC 763

QY 721 TCTTTGGCCAACTCCCTCTCAGATTTGGCAATGCATGACAGAGGAATTTGCTCA 780

Db 764 TCTTTGGCCAACTCCCTCTCAGATTTGGCAATGCATGACAGAGGAATTTGCTCA 823

QY 781 GAAGAAAACATCTATACCATTTGAAGACGTATATGAAGTGAGAGGCCCAATGAT 840

Db 824 GAAGAAAACATCTATACCATTTGAAGACGTATATGAAGTGAGAGGCCCAATGAT 883

QY 841 TATGTCTATGTGACGACGAGGCAACCTCTCAACCTTTGGTTGTGCTTTGCAAG 900

Db 884 TATGTCTATGTGACGACGAGGCAACCTCTCAACCTTTGGTTGTGCTTTGCAAG 943

QY 901 CCA 903

Db 944 CCA 946

RESULT 8

LOCUS ARI43568 2236 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 23 from patent US 6204371.

ACCESSION ARI43568

VERSION ARI43568.1 GI:15104854



KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2236)
TITLE	Levinson, D. Adam.
JOURNAL	Compositions and methods for the treatment and diagnosis of immune disorders
FEATURES	Patent: US 6204371-A 23 20-MAR-2001;
source	Location/Qualifiers 1. .2236 /organism="unknown"
BASE COUNT	586 a 518 c 525 g 606 t 1 others
ORIGIN	
Query Match	100.0%; Score 903; DB 6; Length 2236;
Best Local Similarity	100.0%; Pred. No. 1,2e-277;
Matches 903; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Qy	1 ATGTTTCAACATCTTCCTTTGACTGTGTCTGTCTGTCTGTCTGTCTACTACTACAAAG 60
Db	42 ATGTTTTCACATCTTCCTTTGACTGTGTCTGTCTGTCTGTCTGTCTGTCTACTACTACAAAG 101
Qy	61 TCCCTCAGAGTGGGAAATAGAGCGGAGGCTGGCTCAGAAATGCTTAATCTGCCCCCTTTTAC 120
Db	102 TCCCTCAGAGTGGGAAATAGAGCGGAGGCTGGCTCAGAAATGCTTAATCTGCCCCCTTTTAC 161
Qy	121 ACCCCAGCGCGCCCGAGGGAACCTCGTCCCTCTCTGTGGGCAAGAAGCTGTCTGTG 180
Db	162 ACCCCAGCGCGCCCGAGGGAACCTCGTCCCTCTCTGTGGGCAAGAAGCTGTCTGTG 221
Qy	181 TTGGAATGTGGCAACGTGTGTCTCAGACTGATGAAGAAGGATGTGAATTTATGACATCC 240
Db	222 TTGGAATGTGGCAACGTGTGTCTCAGACTGATGAAGAAGGATGTGAATTTATGACATCC 281
Qy	241 AGATACCTGGCTAAATGGGGATTTCCGCAAAAGGAGATGTGTCCCTGACCAATAGAGATGG 300
Db	282 AGATACCTGGCTAAATGGGGATTTCCGCAAAAGGAGATGTGTCCCTGACCAATAGAGATGG 341
Qy	301 ACTCTAGCAGACAGTGGGATCTACTGTCTGCGGATCCAAATCCAGGCAATATGATGAT 360
Db	342 ACTCTAGCAGACAGTGGGATCTACTGTCTGCGGATCCAAATCCAGGCAATATGATGAT 401
Qy	361 GAAAAATTTAACTTGAAGTTGGTCATCAACCCAGCCAAGTCAACCCCTGACCGACTGTG 420
Db	402 GAAAAATTTAACTTGAAGTTGGTCATCAACCCAGCCAAGTCAACCCCTGACCGACTGTG 461
Qy	421 CAGAGAGCTTCACCTGCAGCCCTTCCAGAGATGCTTACACCAAGGGGACATGGCCACGA 480
Db	462 CAGAGAGCTTCACCTGCAGCCCTTCCAGAGATGCTTACACCAAGGGGACATGGCCACGA 521
Qy	481 GAGACACAGACACTGGGGAGCTCCCTGATATATTAATCTTAACAACAATATCCACATTGGCC 540
Db	522 GAGACACAGACACTGGGGAGCTCCCTGATATATTAATCTTAACAACAATATCCACATTGGCC 581
Qy	541 AATGAGTTACGGGACTTGAATGTGGCCAATGACTTAAGGGACTGTGAGCAACCAATCAGA 600
Db	582 AATGAGTTACGGGACTTGAATGTGGCCAATGACTTAAGGGACTGTGAGCAACCAATCAGA 641
Qy	601 ATAGGCATCTACATGGAGCAGAGGATCTGTGCTGGGCTGGCTCTGTGCTTTATCTTGGCG 660
Db	642 ATAGGCATCTACATGGAGCAGAGGATCTGTGCTGGGCTGGCTCTGTGCTTTATCTTGGCG 701
Qy	661 GCTTTAATTTTCAATGTGATTTCTATAGCAAGAAGAGATACAGAAATTTAAGCCTCATC 720
Db	702 GCTTTAATTTTCAATGTGATTTCTATAGCAAGAAGAGATACAGAAATTTAAGCCTCATC 761
Qy	721 TCTTTTGGCAACCTCCCTCCCTCAGAGATTGGCAAAATGCATGACAGAGGGAATTCGCTCA 780
Db	762 TCTTTTGGCAACCTCCCTCCCTCAGAGATTGGCAAAATGCATGACAGAGGGAATTCGCTCA 821
Qy	781 GAAGAAACATCTATACCATTTGAAGAAAGATGATATGATGTGAGAGAGCCCAATGATAT 840

Db	822	GAAAGAAA	CACTATACCTATTGGAAGAGAAAGAACTATATATGAATGAGAGAGAGCCCAATGAGAT	881
Qy	841	TATTGCTATGT	CAGCAGCAGGCAAGCAACCTCTACAACTTTGGTGTGTGCTTTGCAATG	900
Db	882	TATTGCTATGT	CAGCAGCAGGCAAGCAACCTCTACAACTTTGGTGTGTGCTTTGCAATG	941
Qy	901	CCA	903	
Db	942	CCA	944	
RESULT 9				
LOCUS	AR168937	2236 bp	DNA	linear
DEFINITION	Sequence 23 from patent US 6288218.			
ACCESSION	AR168937			
VERSION	AR168937.1	GI:17905128		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 2236)			
AUTHORS	Levinson D.Adam.			
TITLE	Compositions and methods for the treatment and diagnosis of immune disorders			
JOURNAL	Patent: US 6288218-A 23 11-SEP-2001;			
FEATURES	Location/Qualifiers			
source	1..2236			
BASE COUNT	586 a 518 c 525 g 606 t			1 others
ORIGIN	/organism="unknown"			
Query Match	100.0%;	Score 903;	DB 6;	Length 2236;
Best Local Similarity	100.0%;	Pred. No. 1.2e-277;		
Matches	903;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0.
Qy	1	ATGTTTTCACATCTCTCCCTTTGACTGTGTCTGCTGCTGCTGCTACTACTTACAAGG	60	
Db	42	ATGTTTTCACATCTCTCCCTTTGACTGTGTCTGCTGCTGCTGCTGCTACTACTTACAAGG	101	
Qy	61	TCCTCAGAGTGAATACAGACGCGAGGTGGGTGAGAAATGCTATCTGCGCTCTTAC	120	
Db	102	TCCTCAGAGTGAATACAGACGCGAGGTGGGTGAGAAATGCTATCTGCGCTCTTAC	161	
Qy	121	ACCCAGCGCGCCGAGGAACTCGTGCCTGCTGCTGGGCAAGAGAGCTGTCTGTG	180	
Db	162	ACCCAGCGCGCCGAGGAACTCGTGCCTGCTGCTGGGCAAGAGAGCTGTCTGTG	221	
Qy	181	TTTGAATGTGCAACGTGTGCTCAGACGTATGAAGAGATGTGAATTATTTAGCATCC	240	
Db	222	TTTGAATGTGCAACGTGTGCTCAGACGTATGAAGAGATGTGAATTATTTAGCATCC	281	
Qy	241	AGATACGGCTTAATGGGGATTTCCGAAAGGAATGTGTCCCTGACCATAGAGAAATGTG	300	
Db	282	AGATACGGCTTAATGGGGATTTCCGAAAGGAATGTGTCCCTGACCATAGAGAAATGTG	341	
Qy	301	ACTCTAGAGACAAGTGGGATCTACTGCTGCGGATTCCAAATCCAGAGCATATGATGAT	360	
Db	342	ACTCTAGAGACAAGTGGGATCTACTGCTGCGGATTCCAAATCCAGAGCATATGATGAT	401	
Qy	361	GAATAATTTAACTGAAGTTGGTATCATCAACCAAGCAAGTCAACCCCTGACCGGACTGTG	420	
Db	402	GAATAATTTAACTGAAGTTGGTATCATCAACCAAGCAAGTCAACCCCTGACCGGACTGTG	461	
Qy	421	CAGAGAGCTTCACTGAGCCTTTTCCAAAGATGCTTACCAACAGGGGCAATGGGCCAGCA	480	
Db	462	CAGAGAGCTTCACTGAGCCTTTTCCAAAGATGCTTACCAACAGGGGCAATGGGCCAGCA	521	
Qy	481	GAGACACAGCACTGGGGAGCCTCCCTGATATTAATCTTAACACAAATTTCCATTGGCC	540	
Db	522	GAGACACAGCACTGGGGAGCCTCCCTGATATTAATCTTAACACAAATTTCCATTGGCC	581	
Qy	541	AATGAGTTACGGGACTTGAATTTGGCCAATGACTTACGGGACTCTGAGAACCATCATCAGA	600	





FEATURES		Location/Qualifiers
SOURCE		1..1116
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="2"
		/cell_type="T cell"
gene		1..1116
		/gene="TIM3"
CDS		58..963
		/gene="TIM3"
		/note="TIM-3"
		/codon_start=1
		/product="T cell immunoglobulin mucin-3"
		/protein_id="AA16158.1"
		/db_xref="GI:18182535"
		/translation="MFSLPDPQVLLLLLLLTSSSEVEYRAEVGONAYLPCFYPPA PGNVYPMGKACPFECGNVLTDERDVVWTSRYLNGDFPKGVSLTENVIT ADSGYCCRIQIPGIMDEKFNKLVIPAVTAPTRORPETAAPMLITRGGSP ETQRLGSLPDINLTDISTLANELRBRSLANDLRDSGARTRIGIYGAICAGLALAL FGALIFKYSKKEKIONLSLISLANLPSSGLANVAGINSEENITVEENVYEEVEE PNEYIKVSSRQPSQPLGCRFPAMP"
BASE COUNT	291 a 268 c 271 g 286 t	
ORIGIN		
Query Match	99.8%; Score 901.4; DB 9; Length 1116;	
Best local Similarity	99.9%; Pred. No. 3.5e-277;	
Matches 902; Conservative	0; Mismatches 1; Indels 0; Gaps 0	

QY	661	GCTTTAATTTTCAAAATGTAATTTCTCATAGCAAAAGAAATACAGAAATTTAAGCCTTATC	720
Db	718	GCCTTAATTTTCAAAATGTAATTTCTCATAGCAAAAGAAATACAGAAATTTAAGCCTTATC	777
QY	721	TCCTTGGCAACCCCTCCCTCATAGATTGGCAAAATGCAATAGAGAGGAATTCGGCTCA	780
Db	778	TCCTTGGCAACCCCTCCCTCATAGATTGGCAAAATGCAATAGAGAGGAATTCGGCTCA	837
QY	781	GAAAGAAAACATCTTATACCATTTGAAGAGAGATATGAATGAGAGAGCCCAATGAGTAT	840
Db	838	GAAAGAAAACATCTTATACCATTTGAAGAGAGATATGAATGAGAGAGCCCAATGAGTAT	897
QY	841	TATTGCTATGTACAGACAGAGCAACCTTCACCAACTTTGGGTGTGCTTTGCAATG	900
Db	898	TATTGCTATGTACAGACAGAGCAACCTTCACCAACTTTGGGTGTGCTTTGCAATG	957
QY	901	CCA 903	
Db	958	CCA 960	
RESULT 12			
LOCUS	AF251707	2320 bp	mRNA
DEFINITION	Homo sapiens hepatitis A virus cellular receptor 2 mRNA, complete cds.		
ACCESSION	AF251707		
VERSION	AF251707.1	GI:20330551	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2320)		
AUTHORS	Zhang, W., Wan, T., Li, N. and Cao, X.		
TITLE	Novel human hepatitis A virus cellular receptor		
JOURNAL	2 (bases 1 to 2320)		
REFERENCES	Zhang, W., Wan, T., Li, N. and Cao, X.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-APR-2000) Department of Immunology, Second Military		
JOURNAL	Medical University & Shanghai Brilliance Biotechnology Institute,		
	800 Xiangyin Rd., Shanghai 200433, P.R.China		
FEATURES	Location/Qualifiers		
source	1..2320		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	127..1032		
	/codon_start=1		
	/product="hepatitis A virus cellular receptor 2"		
	/protein_id="PAM19100.1"		
	/db_xref="GI:20330552"		
	/translation="MPSHLPDPDVLILLILLITRSSSEVRYAEVQNAVLPFPYTPAA		
	PGNIVPQMGKAPVECCGVNVLRTDERVNVWTSRYLNGDFRKDVLITENVILL		
	ADSGIVCCRIQIDPGLIMNDEKPNLKIIVPKAKTVPAPLQRPDPAPRMLTTGGHPA		
	ETQGLSGLPDINLIQTISTLANEPLRDSLANLADDSQATITIGIYIGAGICAGLALMI		
	FGALIFRYSKSKTIONLSLNLNIPSGLANAAVAGIRSEBNYITIEENVYEE		
	PNEYIVCVSSKQDSQPLGCRFAMP"		
BASE COUNT	606 a	530 c	554 g 630 t
ORIGIN			
Query Match	99.8%	Score 901.4;	DB 9; Length 2320;
Best Local Similarity	99.9%	Pred. No. 3.9e-277;	
Matches 902; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGTTTACATCTCCCTTGACTGTGTCCTGCTGCTGCTGCTACTACTACCAAG	60
Db	127	ATGTTTACATCTCCCTTGACTGTGTCCTGCTGCTGCTGCTACTACTACCAAG	186
QY	61	TCCTCAGAAGTGAATACAGACGAGAGTGGTCAGAAATCTATCTGCCCTGCTTAC	120
Db	187	TCCTCAGAAGTGAATACAGACGAGAGTGGTCAGAAATCTATCTGCCCTGCTTAC	246



[illegible]

Db	577	GCTGATGAATAA-----	-----GCACTCGAGAAACGAT	605
Qy	598	AGAATAGCATCTATCATCGAGACAGGAGATCTGTGCTGGCTGCGCTCTTATCTTC		657
Db	610	AGAACTGCTATCCACATTTGAGTGAGAGTCTCTGCTGGATTGACCTGGCACTTATCAT		669
Qy	658	GGCGCTTAAATTTTCAATTGTGATTTCTCATGTGCAAGAGAGATACAGAAATTAAGCCTC		717
Db	670	GGTGCTTAAATCTTAAATGATATTTCTGTAGAGAAAGAAATTATCGAGTTTGAAGCCTT		729
Qy	718	ATCTCTTTGGCCAACTCCCTCTCCCTCAGAGATTTGGCAATTTGACAGAGGAAATTCGC		777
Db	730	ATTACACTGGCCAACTTCCCTCTCCAGAGAGGTTGGCAATTCAGAGACAGTCAAGATTCG		789
Qy	778	TCAGAGAAAAATCTATATACCATTTGAGAGAACGTATATGAGTGGAGAGCCCAATGAG		837
Db	790	TCTGAGAAAAATCTACACCATCGAGAGAGAACGTATATGAGTGGAGATTCAAATGAG		849
Qy	838	TATTAATTCCTATATGTCAGACAGCAGCAGAACCTTCACAACTTTGGATTGTGCTTTGCA		897
Db	850	TACTACTGCTATGTCACAGCAGCAGCAGCATCTCGACCGCTCTGAGACTGCCACTTTTAA		909
Qy	898	ATGC		901
Db	910	AGGC		913
RESULT 15				
ARI68930				
LOCUS	ARI68930	2710 bp	DNA	linear
DEFINITION	Sequence 8 from patent US 6288218.			
ACCESSION	ARI68930			
VERSION	ARI68930.1	GI:17905107		
KEYWORDS	unknown.			
SOURCE	unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2710)			
AUTHORS	Levinson, D. Adam.			
TITLE	Compositions and methods for the treatment and diagnosis of immune disorders			
JOURNAL	Patent: US 6288218-A 8 11-SEP-2001;			
FEATURES	Location/Qualifiers			
source	1..2710			
BASE COUNT				
ORIGIN	789 a	592 c	644 g	678 t 7 others
Query Match 50.5%; Score 456; DB 6; Length 2710;				
Best Local Similarity 72.0%; Pred. No. 2.9e-134;				
Matches 651; Conservative 0; Mismatches 220; Indels 33; Gaps 3				
Qy	1	ATGTTTTCACATCTTCCCTTGACTGTGTCTCTGCTGCTGCTGCTGCTGCTACTTACAAAG		60
Db	40	ATGTTTTCAGGCTTAACTCCCTCAACTGTGTCTCTGCTGCTGCTGCTGCTGCTACTTCCAAAG		99
Qy	61	TCCTCAGAAATGGAATACAGACGCGAGGTCGCTGCAATGCTACTGTCGCCCTTCTAC		120
Db	100	TCATTTGAGAAATGTTATTAAGTTTGAGTTGATTAATAATGCTTATCTTGCCCTGCACTTAC		159
Qy	121	ACCCAGCCGCCCCAGGGAACCTCTGTGCTCTGCTGCGGCAAGAGACCTGTCTGTG		180
Db	160	ACTCTACTACATCTGAGGACACTTGTGCTATGTGCTGTGGGCAAGGATTTGTCCTTG		219
Qy	181	TTTGAATGTGGCAACGTGTGCTCAGACCTGATGAAGAGATGTGAATTAATTGACATC-		239
Db	220	TCACAGTGTACCAATGAATTTCTCAGAACTATGAAAGAAATGTGACATATCAGAAATCC		279
Qy	240	--CAGATCTGCTTAAATGAGGGAATTTCCGCAAGAGAGATGTGCTCGACCATAGAGAT		297
Db	280	AGCAGATTCACAGCTTAAAGGCGGATCTTCAACAAAGAGATGTGTCTTATCATTAAGAAAT		339
Qy	298	GTGACTCTTAGACAGACAGTGGGATCTACTGTGCTCGCGATCCAAATCCAGCATTAATGAT		357

```
Db 340 G|T|A|C|T|G|G|A|T|A|C|C|A|T|G|G|A|C|C|T|A|C|T|G|T|G|C|A|G|G|A|T|A|C|A|G|T|T|C|C|T|G|T|T|A|T|A|T| 399
QY 358 G|A|T|G|A|A|A|A|T|T|T|A|C|C|T|G|A|G|T|G|T|C|A|T|C|A|A|C|A|G|C|C|A|G|G|T|C|A|C|C|C|T|G|C|A|C|C|A|C|T| 417
Db 400 G|A|T|A|A|A|A|A|T|T|A|A|C|T|G|A|A|T|T|A|G|A|C|A|T|C|A|A|G|C|A|G|C|C|A|G|G|T|C|A|C|C|A|C|T|C|A|G|A|C|T| 459
QY 418 C|T|G|C|A|G|A|G|A|C|T|T|C|A|C|T|G|C|A|G|C|T|T|C|C|A|G|A|T|G|T|T|C|C|A|G|G|G|A|C|A|T|G|G|C|C|A| 477
Db 460 G|C|C|C|A|T|G|G|G|A|C|T|C|T|A|C|T|A|C|A|G|C|T|T|C|C|A|G|A|A|C|C|T|A|C|A|C|G|G|A|G|A|A|T|G|---T 516
QY 478 G|C|A|G|A|G|A|C|A|G|A|C|A|C|T|G|G|G|A|G|C|T|C|C|G|A|T|A|T|A|A|T|C|T|A|C|A|C|A|A|T|A|T|C|C|A|C|A|T|G| 537
Db 517 T|C|A|G|A|G|A|C|A|G|A|C|A|C|T|G|T|G|A|C|C|T|C|A|T|A|T|A|C|A|T|G|A|A|C|A|A|A|A|T|T|C|C|A|C|A|T|G| 576
QY 538 G|C|C|A|T|G|A|G|T|T|A|C|G|G|G|A|C|T|C|T|A|G|A|T|G|C|C|A|T|G|C|T|T|A|C|G|G|A|C|T|T|G|A|G|C|A|C|C|A|T|C| 597
Db 577 G|C|T|G|A|T|G|A|A|T|T|A|-----G|G|A|C|T|C|T|G|G|A|A|A|C|G|A|T|C| 609
QY 598 A|G|A|T|A|G|G|C|A|T|T|A|C|A|T|G|G|A|G|G|G|A|T|T|G|T|G|T|G|G|G|C|T|G|G|C|T|G|G|C|T|T|A|T|C|T|T|C| 657
Db 610 A|G|A|C|T|G|C|T|A|T|C|C|A|C|A|T|T|G|G|A|G|G|G|A|G|T|C|T|G|T|G|G|G|T|T|G|A|C|C|T|G|G|C|A|C|T|T|A|T|C|A|T| 669
QY 658 G|G|G|C|T|T|A|T|T|T|C|A|A|T|G|T|A|T|T|C|T|C|A|T|A|G|C|A|A|G|A|A|G|A|T|A|C|A|G|A|T|T|A|A|G|C|T|C| 717
Db 670 G|G|T|G|T|T|A|T|C|C|T|T|A|A|T|G|T|A|T|T|C|T|G|T|A|G|A|A|A|A|A|G|T|A|T|G|A|G|C|T|T| 729
QY 718 A|T|C|T|T|T|G|C|C|A|C|C|T|C|C|C|C|T|C|A|G|A|T|T|G|G|C|A|A|T|G|C|A|G|T|A|G|C|A|G|A|G|G|A|T|T|C|G|C| 777
Db 730 A|T|T|A|C|A|C|T|G|G|C|C|A|C|T|G|C|C|C|A|G|G|G|T|T|G|G|C|A|A|T|G|C|A|G|A|G|C|A|G|T|C|A|G|A|T|T|C|G|C| 789
QY 778 T|C|A|G|A|G|A|A|A|C|A|T|T|A|T|A|C|A|T|T|G|A|A|G|A|A|C|G|T|A|T|A|T|G|A|G|G|A|G|G|A|G|C|C|A|T|G|A|G| 837
Db 790 T|C|G|A|G|A|A|A|A|T|A|T|C|T|A|C|A|C|C|A|T|C|A|G|A|G|A|A|C|G|T|A|T|A|T|G|A|G|A|T|T|C|A|A|A|T|G|A|G| 849
QY 838 T|A|T|A|T|G|C|T|A|T|G|T|C|A|G|C|A|G|C|A|G|C|A|C|C|T|C|A|C|A|C|C|T|T|G|G|T|T|G|C|T|T|G|C|A| 897
Db 850 T|A|C|T|A|C|T|G|C|T|A|G|T|C|A|G|C|A|G|C|A|G|C|A|T|C|T|G|A|C|C|G|C|C|T|G|A|C|T|G|C|C|A|C|T|T|T|A| 909
QY 898 A|T|G|C| 901
Db 910 A|G|G|C| 913
```

Search completed: November 21, 2003, 23:34:31  
Job time : 3377.32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:32:52 ; Search time 248.931 Seconds  
(without alignments)  
9792.239 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903  
Sequence: 1 atgtttccatccatccctt.....gttgcgttcgatcgcca 903

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq.19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	903	100.0	903	21	AAAS1905
2	903	100.0	903	22	AAI70281
3	903	100.0	903	22	AAF82629
4	903	100.0	903	22	AAAC92155
5	903	100.0	903	22	AAAF23478
6	903	100.0	903	25	ABQ77041
7	903	100.0	1116	25	ABZ68338
8	903	100.0	1756	22	AAAS22714

9	903	100.0	1772	22	AAH15441
10	903	100.0	2236	21	AAAS1904
11	903	100.0	2236	22	AAI70254
12	903	100.0	2236	22	AAF82616
13	903	100.0	2236	22	AAAC90993
14	903	100.0	2236	22	AAAF23465
15	903	100.0	2236	22	ABSS5322
16	903	100.0	2236	25	ABQ77040
17	903	100.0	2237	17	AAI78267
18	902.6	100.0	1475	24	ABL90805
19	901.4	99.8	1116	25	ABZ68339
20	901.4	99.8	1800	22	AAAS22478
21	799	88.5	1326	25	AAI73381
22	712	78.8	769	22	AAH04055
23	585	64.8	2091	22	AAK35717
24	522.8	57.9	555	20	AAK35716
25	456	50.5	2710	21	AAAS1898
26	456	50.5	2710	22	AAAI70263
27	456	50.5	2710	22	AAAF82609
28	456	50.5	2710	22	AAAC90986
29	456	50.5	2710	22	AAAF23458
30	456	50.5	2710	24	ABSS5316
31	456	50.5	2710	25	ABQ77037
32	456	50.5	2712	17	AAI78265
33	456	50.5	2725	25	ABZ68328
34	455	50.4	529	20	AAV87937
35	442.6	49.0	843	21	AAAS1918
36	442.6	49.0	843	22	AAI70255
37	442.6	49.0	843	22	AAF82615
38	442.6	49.0	843	22	AAAC90992
39	442.6	49.0	843	22	AAAF23464
40	442.6	49.0	843	25	ABQ77038
41	440.8	48.8	862	25	ABZ68329
42	440.8	48.9	1203	21	AAAB8799
43	374.2	41.4	438	20	AAAX4096
44	74	8.2	4209	25	AAAC50985
45	72.6	8.0	1079	25	ABZ68335

#### ALIGNMENTS

RESULT 1	AAAS1905	standard; cDNA, 903 BP.
ID	AAAS1905	
AC	AAAS1905;	
XX		
DT	31-OCT-2000	(first entry)
DE	Human T helper cell differentially expressed gene 200 cDNA.	
XX		
KW	T helper cell; differential expression; 200 gene; immunomodulator;	
KW	anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;	
KW	thymomimetic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;	
KW	protozoacide; lymphocyte; modulator; gene therapy; ss.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..903
FT		/tag= a
FT		/partial
FT	sig_peptide	1..60
FT		/tag= b
FT	mat_peptide	61..903
FT		/tag= c
XX		
PN	US6084083-A.	
XX		
PD	04-JUL-2000.	
XX		
PF	28-MAR-1997;	97US-0829525.



PS Disclosure; Column 135-136; 108pp; English.

The present sequence is that of cDNA comprising the complete open reading frame of the human 200 gene. The encoded protein (see *AAIM0223*) is a receptor of the Ig superfamily class. The cDNA was isolated from a human lymphocyte cDNA library using mouse 200 gene cDNA as probe. The nucleotide sequence contained within B. coli clone felt 200C is deposited as ATCC 69967. The human 200 gene is preferentially expressed in mature, fully differentiated T helper subpopulation TH1 cells relative to subpopulation TH2 cells. The gene can be used diagnostically or as a target for therapeutic intervention for the treatment of immune disorders. A claimed method for diagnosing a TH cell subpopulation-related immune disorder involves detecting the level of a 200 gene product, or an RNA encoding it, so that if the level differs from that in a control sample, the disorder is diagnosed. The disorder is especially a TH1 cell subpopulation-related immune disorder, such as multiple sclerosis, psoriasis or insulin-dependent diabetes (claimed). In addition to the 200 gene, the invention provides other genes that are differentially expressed within and among TH cells and TH cell subpopulations and which can be used in methods for the diagnosis, prognosis, evaluation and treatment of TH cell subpopulation-related disorders, for the identification of subjects exhibiting a predisposition to such conditions, for monitoring patients undergoing clinical evaluation for the treatment of such disorders, and for monitoring the efficacy of compounds used in clinical trials. Other immune disorders that can be treated/diagnosed include Crohn's disease, reactive arthritis, Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis, Grave's disease, contact dermatitis, graft rejection, graft versus host disease, sarcoidosis, atopic conditions, asthma, allergy, allergic rhinitis, food allergy, eosinophilia, conjunctivitis, glomerular nephritis, helminthic infection (e.g. leishmaniasis), viral infection (e.g. HIV), and bacterial infection (e.g. tuberculosis and lepromatous leprosy).

SQ Sequence 903 BP; 242 A; 225 C; 221 G; 215 T; 0 other;

Query Match	Score	DB	Length
100.0%	903	22	903

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGTTTCAACATTTTCCCTTTGACTGTGCTGCTGCTGCTGCTACACTTACAAAG	60
Db	1	ATGTTTCAACATTTTCCCTTTGACTGTGCTGCTGCTGCTGCTACACTTACAAAG	60
Qy	61	TCTCTCAAAAGTGGATATACAGACCGGAGGTGGTGCAGATTCCTATCTGCCCCCTTCTAC	120
Db	61	TCTCTCAAAAGTGGATATACAGACCGGAGGTGGTGCAGATTCCTATCTGCCCCCTTCTAC	120
Qy	121	ACCCGACGCGCCCCGAGGGAACCTCGTCCCGCTCTGCTGGGCAAGAGCCGTGCTGTG	180
Db	121	ACCCGACGCGCCCCGAGGGAACCTCGTCCCGCTCTGCTGGGCAAGAGCCGTGCTGTG	180
Qy	181	TTTGAATGTGGCAACCTGTGTCTACAGACTGATGAAAGGAGTGTGAATTTATTTGACATCC	240
Db	181	TTTGAATGTGGCAACCTGTGTCTACAGACTGATGAAAGGAGTGTGAATTTATTTGACATCC	240
Qy	241	AGATACTGGCTAAATGGGGATTTCCGCCAAGAGATGTGTCTCTGACCAATAGAGATGTG	300
Db	241	AGATACTGGCTAAATGGGGATTTCCGCCAAGAGATGTGTCTCTGACCAATAGAGATGTG	300
Qy	301	ACCTTAGCAGACGATGGGATCTACTGTGCGGGATTCGAAATCCGACGGCATATGAT	360
Db	301	ACCTTAGCAGACGATGGGATCTACTGTGCGGGATTCGAAATCCGACGGCATATGAT	360
Qy	361	GAATAATTTAACTGAAGTTGTGATCAAAACGACCAAGGTACCCCTTGACACGACTCTG	420
Db	361	GAATAATTTAACTGAAGTTGTGATCAAAACGACCAAGGTACCCCTTGACACGACTCTG	420
Qy	421	CAGAGAGACTTCACTGACGCTTTCCAAAGAGTCTTACACAGGGGACATGGCCACGA	480
Db	421	CAGAGAGACTTCACTGACGCTTTCCAAAGAGTCTTACACAGGGGACATGGCCACGA	480

QY	481	GAGCACA	GACACTG	GGAGG	AGCTCCCTG	ATATTAATCTAA	CAACAATATCC	CAATGGCC	540	
Db	481	GAGCACA	GACACTG	GGGAG	AGCCCTCCCTG	ATATTAATCTAA	CAACAATATCC	CAATGGCC	540	
QY	541	AATGAGTTA	CGGGACTCT	TAAGTTGG	CCAAATGA	CTTACGGGACTCT	GGAGCAACCAT	CAGA	600	
Db	541	AATGAGTTA	CGGGAGCTCT	TAGATTGG	CCAAATGA	CTTACGGGACTCT	GGAGCAACCAT	CAGA	600	
QY	601	ATAGGCAT	CTACATCG	AGAGGAGAT	CTGTGCTGGG	CTGGCTCTG	AGCTTAACTTC	GGG	660	
Db	601	ATAGGCAT	CTACATCG	AGGAGAT	CTGTGCTGGG	CTGGCTCTTAA	CTTCGGC		660	
QY	661	GCTTTAATTTT	CAATG	GTATTTCT	CATAGCA	AGAGAGATAC	CAGAAATTTAA	GCCCTCATC	720	
Db	661	GCTTTAATTTT	CAATG	TATTTCT	CATAGCA	AGAGAGATAC	CAGAAATTTAA	GCCCTCATC	720	
QY	721	TCTTTGGC	CAACCTCC	CTCCCTC	CAGAGATTGG	CAAAATGC	AGTGC	AGAGGAAATTC	GCTCA	780
Db	721	TCTTTGGC	CAACCTCC	CTCCCTC	CAGAGATTGG	CAAAATGC	AGTGC	AGAGGAAATTC	GCTCA	780
QY	781	GAGAAAA	CATCTAT	ATACATTGA	AGAGACGTATAT	GAAAGTGG	AGAGGCCCAT	GATGAT	840	
Db	781	GAGAAAA	CATCTAT	ATACATTGA	AGAGACGTATAT	GAAAGTGG	AGAGGCCCAT	GATGAT	840	
QY	841	TATTGCTAT	GTTCAG	CAGCAGG	CAGGCAAGCA	CCCTTC	CAACCTTTGG	GTTCGCTTTG	CAATG	900
Db	841	TATTGCTAT	GTTCAG	CAGCAGG	CAGGCAAGCA	CCCTTC	CAACCTTTGG	GTTCGCTTTG	CAATG	900
QY	901	CCA	903							
Db	901	CCA	903							

### RESULT 3

ID AAF82629 standard; cDNA; 903 BP.

AC AAF82629

DT 18-JUN-2001 (first entry)

DE Human TH1 specific 200 gene open reading frame

KW Human; T helper cell; TH1; TH2; immunomodulator;

KW antibacterial; T helper lymphocyte modulator; gene therapy

KW infection; ss.

OS Homo sapiens.

**FH Key**

13

2000  
 2001  
 2002  
 2003  
 2004  
 2005  
 2006  
 2007  
 2008  
 2009  
 2010  
 2011  
 2012  
 2013  
 2014  
 2015  
 2016  
 2017  
 2018  
 2019  
 2020  
 2021  
 2022  
 2023  
 2024  
 2025  
 2026  
 2027  
 2028  
 2029  
 2030  
 2031  
 2032  
 2033  
 2034  
 2035  
 2036  
 2037  
 2038  
 2039  
 2040  
 2041  
 2042  
 2043  
 2044  
 2045  
 2046  
 2047  
 2048  
 2049  
 2050  
 2051  
 2052  
 2053  
 2054  
 2055  
 2056  
 2057  
 2058  
 2059  
 2060  
 2061  
 2062  
 2063  
 2064  
 2065  
 2066  
 2067  
 2068  
 2069  
 2070  
 2071  
 2072  
 2073  
 2074  
 2075  
 2076  
 2077  
 2078  
 2079  
 2080  
 2081  
 2082  
 2083  
 2084  
 2085  
 2086  
 2087  
 2088  
 2089  
 2090  
 2091  
 2092  
 2093  
 2094  
 2095  
 2096  
 2097  
 2098  
 2099  
 2100  
 2101  
 2102  
 2103  
 2104  
 2105  
 2106  
 2107  
 2108  
 2109  
 2110  
 2111  
 2112  
 2113  
 2114  
 2115  
 2116  
 2117  
 2118  
 2119  
 2120  
 2121  
 2122  
 2123  
 2124  
 2125  
 2126  
 2127  
 2128  
 2129  
 2130  
 2131  
 2132  
 2133  
 2134  
 2135  
 2136  
 2137  
 2138  
 2139  
 2140  
 2141  
 2142  
 2143  
 2144  
 2145  
 2146  
 2147  
 2148  
 2149  
 2150  
 2151  
 2152  
 2153  
 2154  
 2155  
 2156  
 2157  
 2158  
 2159  
 2160  
 2161  
 2162  
 2163  
 2164  
 2165  
 2166  
 2167  
 2168  
 2169  
 2170  
 2171  
 2172  
 2173  
 2174  
 2175  
 2176  
 2177  
 2178  
 2179  
 2180  
 2181  
 2182  
 2183  
 2184  
 2185  
 2186  
 2187  
 2188  
 2189  
 2190  
 2191  
 2192  
 2193  
 2194  
 2195  
 2196  
 2197  
 2198  
 2199  
 2200  
 2201  
 2202  
 2203  
 2204  
 2205  
 2206  
 2207  
 2208  
 2209  
 2210  
 2211  
 2212  
 2213  
 2214  
 2215  
 2216  
 2217  
 2218  
 2219  
 2220  
 2221  
 2222  
 2223  
 2224  
 2225  
 2226  
 2227  
 2228  
 2229  
 2230  
 2231  
 2232  
 2233  
 2234  
 2235  
 2236  
 2237  
 2238  
 2239  
 2240  
 2241  
 2242  
 2243  
 2244  
 2245  
 2246  
 2247  
 2248  
 2249  
 2250  
 2251  
 2252  
 2253  
 2254  
 2255  
 2256  
 2257  
 2258  
 2259  
 2260  
 2261  
 2262  
 2263  
 2264  
 2265  
 2266  
 2267  
 2268  
 2269  
 2270  
 2271  
 2272  
 2273  
 2274  
 2275  
 2276  
 2277  
 2278  
 2279  
 2280  
 2281  
 2282  
 2283  
 2284  
 2285  
 2286  
 2287  
 2288  
 2289  
 2290  
 2291  
 2292  
 2293  
 2294  
 2295  
 2296  
 2297  
 2298  
 2299  
 2300  
 2301  
 2302  
 2303  
 2304  
 2305  
 2306  
 2307  
 2308  
 2309  
 2310  
 2311  
 2312  
 2313  
 2314  
 2315  
 2316  
 2317  
 2318  
 2319  
 2320  
 2321  
 2322  
 2323  
 2324  
 2325  
 2326  
 2327  
 2328  
 2329  
 2330  
 2331  
 2332  
 2333  
 2334  
 2335  
 2336  
 2337  
 2338  
 2339  
 2340  
 2341  
 2342  
 2343  
 2344  
 2345  
 2346  
 2347  
 2348  
 2349  
 2350  
 2351  
 2352  
 2353  
 2354  
 2355  
 2356  
 2357  
 2358  
 2359  
 2360  
 2361  
 2362  
 2363  
 2364  
 2365  
 2366  
 2367  
 2368  
 2369  
 2370  
 2371  
 2372  
 2373  
 2374  
 2375  
 2376  
 2377  
 2378  
 2379  
 2380  
 2381  
 2382  
 2383  
 2384  
 2385  
 2386  
 2387  
 2388  
 2389  
 2390  
 2391  
 2392  
 2393  
 2394  
 2395  
 2396  
 2397  
 2398  
 2399  
 2400  
 2401  
 2402  
 2403  
 2404  
 2405  
 2406  
 2407  
 2408  
 2409  
 2410  
 2411  
 2412  
 2413  
 2414  
 2415  
 2416  
 2417  
 2418  
 2419  
 2420  
 2421  
 2422  
 2423  
 2424  
 2425  
 2426  
 2427  
 2428  
 2429  
 2430  
 2431  
 2432  
 2433  
 2434  
 2435  
 2436  
 2437  
 2438  
 2439  
 2440  
 2441  
 2442  
 2443  
 2444  
 2445  
 2446  
 2447  
 2448  
 2449  
 2450  
 2451  
 2452  
 2453  
 2454

•

2  
3  
4  
5  
6  
7  
8

XX  
3  
100

PR 07-JUN-1995;  
XX

PA (MILL-) MILLENNIUM PHARM INC.

PI Levinson DA,

DR WPI; 2001-272703/28











Qy	361	GAATAAATTAACTGAAGTTGGTCATCAAAACGAGCAAGATGACCCCTGGCACCACTGTG	420
Dy	361	GAATAAATTAACTGAAGTTGGTCATCAAAACGAGCAAGATGACCCCTGGCACCACTGTG	420
Qy	421	CAGAGGACCTTCACTGACGCTTTCCAGAGATGCTTACCAACAGGGGACATGCGCCAGCA	480
Dy	421	CAGAGGACCTTCACTGACGCTTTCCAGAGATGCTTACCAACAGGGGACATGCGCCAGCA	480
Qy	481	GAGACACAGACACTGGGAGACCTCCCTGATATAATCTTAAACAATAATTCACATTTGAGCC	540
Dy	481	GAGACACAGACACTGGGAGACCTCCCTGATATAATCTTAAACAATAATTCACATTTGAGCC	540
Qy	541	AATGAGTTACGGGACTCTGATTTGGCCAAATGACTTACGGGACTCTGAGCAACCATCAGA	600
Dy	541	AATGAGTTACGGGACTCTGATTTGGCCAAATGACTTACGGGACTCTGAGCAACCATCAGA	600
Qy	601	ATAGGCATCTACATCCGAGCAGGGAATCTGTGCTGGGCTGGCTGTGCTTTATCTTGGGC	660
Dy	601	ATAGGCATCTACATCCGAGCAGGGAATCTGTGCTGGGCTGGCTGTGCTTTATCTTGGGC	660
Qy	661	GCTTTAATTTTCAATGATGATTCATCTACAGAGAGAAATPACAGATTTAAGCTCATC	720
Dy	661	GCTTTAATTTTCAATGATGATTCATCTACAGAGAGAAATPACAGATTTAAGCTCATC	720
Qy	721	TCTTTGGCCAACTCCCTCCCTCAGAGATTGGCAAAATGACAGTACAGAGGGAATTCGCTCA	780
Dy	721	TCTTTGGCCAACTCCCTCCCTCAGAGATTGGCAAAATGACAGTACAGAGGGAATTCGCTCA	780
Qy	781	GAAGAAAAACATCTATACCATTTGAAGAAACGATATGAAAGTGGAGAGACCCAAATGAGTAT	840
Dy	781	GAAGAAAAACATCTATACCATTTGAAGAAACGATATGAAAGTGGAGAGACCCAAATGAGTAT	840
Qy	841	TATTCATATGTCAGACGAGGAGCAACCTTCAACAACCTTTGGGTTGTGTGCTTTCAGATG	900
Dy	841	TATTCATATGTCAGACGAGGAGCAACCTTCAACAACCTTTGGGTTGTGTGCTTTCAGATG	900
Qy	901	CCA 903	
Dy	901	CCA 903	
RESULT 6			
ABQ7041			
ID	ABQ7041	standard; cDNA; 903 BP.	
AC	ABQ7041;		
XX			
DT	09-APR-2003	(first entry)	
XX			
DE		Human TH1-associated 200 gene SEQ ID 37.	
XX			
KW		T-cell receptor; TH; T helper cell; gene; 103 gene; TH2 cell; human;	
KW		TH2 cell marker; TH1 cell; protozoacide; antibacterial; virucide;	
KW		immunosuppressive; antiinflammatory; antiarthritis; antidiabetic;	
KW		neuroprotective; dermatological; antihypoid; antipsoriatic; helminthic;	
KW		nephrotoxic; asthmatic antiallergic; CD8 agonist; CD4 agonist;	
KW		interleukin agonist; bacterial; viral infection; immune disorder;	
KW		Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;	
KW		Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;	
KW		graft rejection; graft versus host disease; asthma; glomerulonephritis;	
KW		allergy; gene therapy; TH cell subpopulation; ss.	
OS		Homo sapiens.	
XX			
PN		US6455685-B1.	
XX			
PD		24-SEP-2002.	
XX			
PF		27-FEB-1998; 98US-0032337.	
XX			
PR		03-MAR-1995; 95US-0398663.	
PR		07-JUN-1995; 95US-0487748.	

PR 01-MAR-1996; 96US-0609583.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Levinson DA;  
XX  
XX WPI; 2003-066247/06.  
DR P-PSDB; ABG73718.  
XX  
PT Identifying a test compound that binds to a 103 gene product or is a  
PT TH2 cell marker, useful for the identification and therapeutic use of  
PT compounds as treatments of helminthic, bacterial and viral infections,  
XX -and immune disorders -  
XX  
PS Example 9; Figure 24A-C; 131pp; English.

This invention describes a novel method for identifying a test compound that binds to a 103 gene product or is a TH2 cell marker. The method comprises contacting a test compound with an immobilised 103 gene product, removing unbound test compound or separating the complex from the reaction mixture, and detecting the complex. Identifying a test compound that binds to a 103 gene product alternatively comprises: (a) contacting a test compound with a cell engineered to express a 103 gene product or co-expressing a 103 gene product and a test compound in a cell; and (b) removing unbound 103 gene product. Identifying a test compound that is a TH2 cell marker further comprises detecting a complex, where contacting the test compound with a TH2 and TH1 cell for the test compound to bind either cell and detecting binding where it indicates the test compound as a TH2 cell marker. The products described in the invention have protozoacide, antibacterial, antiviral, immunosuppressive, antiinflammatory, antiallergic, antidiabetic, neuroprotective, dermatological, antihypoid, antipsoriatic, nephrotoxic, antiaslathmic; and antiallergic activity and can act as CD8, CD4 and interleukin agonists. The methods and compositions of the present invention are useful for the identification and therapeutic use of compounds as treatments of helminthic, bacterial and viral infections and immune disorders such as Crohn's disease, reactive arthritis, diabetes, multiple sclerosis, Hashimoto's thyroiditis, Graves' disease, contact dermatitis, psoriasis, graft rejection, graft versus host disease, asthma, allergy and glomerulonephritis. They can also be used for gene therapy, for the diagnostic evaluation and prognosis of TH cell subpopulation-related disorders, identification of subjects exhibiting a predisposition to such conditions, monitoring undergoing clinical evaluation and efficacy for the treatment of the disorders. This sequence represents a T-Helper cell associated polynucleotide described in the disclosure of the invention.

**SQ** Sequence 903 BP; 242 A; 225 C; 221 G; 215 T; 0 other;

Query Match	Score	DB	Length
100.0%	903	25	903

Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y 1 ATGTTTCAATCTTCCCTTGACTGTCTCTGCTGCTACTACTTACAAGG 60

D5 1 ATGTTTCACATCTTCCCTTGGACTGTGTCCCTGCTGCTGCTGCTACTACTACAAG 60

61 TCCTCAGAGTGGATAACAGAGCGAGGTCGGTCAGATGCCATCTGCCCCGTCCTAC 120

Db 61 TCCTCAGAGTGGATAACAGAGCCGAGGTCGGTCAGAAAGCCATCTGCCCCCTCTCTAC 120

121 ACCCCAGCCGCCCCAGGGAACTCTGTCGCCCTCTCTGCGGCAAGGAGCCCTGTCTCTGTG 180

DB 121 ACCCCAGCCGCCCAGGACCTCGTGCCTGCTGGGCAATGAGCTGTCCTG 180

181 T T G A A T G T G C A A C T G G T G C T C A G G A C T G A T G A A A G G G A T G T A A T A T T G G A C A T C C 240

DU 101-111GHH!G1G9C!HAC!G1G9C!CHABAC!GH!GHH!BDBAH!B1B1BH!BGBCH!CC 270

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 104

[illegible]

QY		30	ACTCTGACGACAGTGGGATCTCACTGTCGCCGGATCCAAATCCAGGCATTAAAGAAT	360			
Dd		301	ACTTGCGACAGCAGTGGATCTACTGTCTGCCGATCCAAATCCAGGCATTAAAGAAT	360			
QY		361	GAAAAATTTAACCCTGAAGTTGTCTCAATCAAACGACCAAGCTCACCCCTGCACGACTCTG	420			
Dd		361	GAAAATTTTAACCTGAAGTTGTCTCAATCAAACGACCAAGCTCACCCCTGCACGACTCTG	420			
QY		421	CAGAGAGCTTCACTGCGAGCTTTTCCAGAGATGCTTACACCAAGGGGACATGCCCCAGCA	480			
Dd		421	CAGAGAGACTTCACTGCGAGCTTTTCCAGAGATGCTTACACCAAGGGGACATGCCCCAGCA	480			
QY		481	GAGACACAGACACTGGGGAGCTCCCTCATATAATTAATCTAACAAATATOCATTGGCC	540			
Dd		481	GAGACACAGACACTGGGGAGCTCCCTCATATAATTAATCTAACAAATATOCATTGGCC	540			
QY		541	AATGAGTTACGGGACTCTAGATTGGCCAATGACTTACGGGACTCTGGAGCAACCATGACA	600			
Dd		541	AATGAGTTACGGGACTCTAGATTGGCCAATGACTTACGGGACTCTGGAGCAACCATGACA	600			
QY		601	ATAGGCATCTACATCGGACGAGGATCTGTGCTGGGCTGGCTCTTATCTTGCGC	660			
Dd		601	ATAGGCATCTACATCGGACGAGGATCTGTGCTGGGCTGGCTCTTATCTTGCGC	660			
QY		661	GCTTTAATTTTCAAATGTTATCTCATAGCAAAAGAAAGATACAGAAATTTAAGCTCATC	720			
Dd		661	GCTTTAATTTTCAAATGTTATCTCATAGCAAAAGAAAGATACAGAAATTTAAGCTCATC	720			
QY		721	TCTTTGGCCAACTCCCTCCCTCAGGATTGGCAATGAGTAGAGGGAATTCGGTCA	780			
Dd		721	TCTTTGGCCAACTCCCTCCCTCAGGATTGGCAATGAGTAGAGGGAATTCGGTCA	780			
QY		781	GAGAAAAACATCTATACCAATTGAAGAGAACGTATATGAATGAGAGAGGCCAATGAGTAT	840			
Dd		781	GAGAAAAACATCTATACCAATTGAAGAGAACGTATATGAATGAGAGAGGCCAATGAGTAT	840			
QY		841	TATTGCTATGTACGACGACGAGCAACCCCTCACACTTTGGGTGTGCTTTGCAATG	900			
Dd		841	TATTGCTATGTACGACGACGAGCAACCCCTCACACTTTGGGTGTGCTTTGCAATG	900			
QY		901	CCA 903				
Dd		901	CCA 903				
RESULT 7							
ABZ68338							
ID	ABZ68338	standard; DNA; 1116 BP.					
XX							
AC	ABZ68338;						
XX							
DT	22-Apr-2003	(first entry)					
DE	Nucleotide sequence of human TIM-3 allele 1.						
KX							
KW	T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;						
KW	TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;						
KW	myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;						
KW	allergic T cell response; autoimmune disease; gene; ss.						
OS							
XX	Homo sapiens.						
XX							
FH	Key	Location/Qualifiers					
FT	CDS	58..963					
FT		/*tag= a					
FT		/product= "TIM-3"					
XX		/transl_except= (pos:475..477,aa:Arg)					
XX							
XX	MO2003002722-A2.						
XX							
XX	09-JAN-2003.						
XX							
PF	01-JUL-2002; 2002WC-US20890.						

XX 29-JUN-2001, 2001US-302344P.  
 XR  
 PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PI  
 PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
 XX  
 DR WPI: 2003-210268/20.  
 DR P-PsDE, ABP70444.  
 XX  
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 PS  
 PS Claim 5, Page 89-90; 94pp; English.  
 XX  
 CC The present sequence encodes a human T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1, TIM-2,  
 CC TIM-3, and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved IgY and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperreactivity and allergic T cell responses, and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for  
 CC treating cancer, asthma, allergies, eczema or autoimmune diseases.  
 XX  
 QD Sequence 1116 BP; 291 A; 268 C; 270 G; 287 T; 0 other;

Query Match	Best Local Similarity	100.0%	Score 903;	DB 25;	Length 1116;				
Matches	903;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	ATGTTTTACATCTCTCCCTTTGACTGTGTCTGTCTGTCTGTCTACTACTTAACAAGG	60						
Db	58	ATGTTTTACATCTTCTCCCTTTGACTGTGTCTGTCTGTCTGTCTACTACTTAACAAGG	117						
Qy	61	TCCTCAGAGTGGAAATACAGAGCGGAGGTGCGTCAGAATGCTTATCTGCCCTGCTTCTAC	120						
Db	118	TCCTCAGAGTGGAAATACAGAGCGGAGGTGCGTCAGAATGCTTATCTGCCCTGCTTCTAC	177						
Qy	121	ACCCAGCGCGCCAGGGAACTGTGCGCGCTGTCTGGGGCAAGAGCGCTGTCTGTG	180						
Db	178	ACCCAGCGCGCCAGGGAACTGTGCGCGCTGTCTGTCTGGGGCAAGAGCGCTGTCTGTG	237						
Qy	181	TTTGAATGTGGCAAGTGTGCTCAGGACTGATAAAGGAGTGTGAATTATTTGACATCC	240						
Db	238	TTTGAATGTGGCAAGTGTGCTCAGGACTGATAAAGGAGTGTGAATTATTTGACATCC	297						
Qy	241	AGATACTGGCTAAATGGGGATTTTCCGCAAGAGAGATGTCTCCCTGACCATAGAGATGTG	300						
Db	298	AGATACTGGCTAAATGGGGATTTTCCGCAAGAGAGATGTCTCCCTGACCATAGAGATGTG	357						
Qy	301	ACTTAGAGACAATGGGATCTACTGTCGCGCCGGATCCAAATCCAGAGATATGAATAT	360						
Db	358	ACTTAGAGACAATGGGATCTACTGTCGCGCGATCCAAATCCAGAGATATGAATAT	417						
Qy	361	GAATAATTTAATCTGAAGTGTGTCATCAACACAGCAAGTCAACCCCTGCACCGACTGTG	420						
Db	418	GAATAATTTAATCTGAAGTGTGTCATCAACACAGCAAGTCAACCCCTGCACCGACTGTG	477						
Qy	421	CAGAGAGCTTCACTGCAGCCCTTTCCAAAGATGCTTTACACACAGGGACATGCGCCACGA	480						
Db	478	CAGAGAGCTTCACTGCAGCCCTTTCCAAAGATGCTTTACACACAGGGACATGCGCCACGA	537						
Qy	481	GAGACACAGACACTGGGGAGCCTCCCTGTATTTAAATCTAACCAAAATTTCCATTTGGCC	540						
Db	538	GAGACACAGACACTGGGGAGCCTCCCTGTATTTAAATCTAACCAAAATTTCCATTTGGCC	597						

QY 541 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 600  
 DB 538 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 657  
 QY 601 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 660  
 DB 658 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 717  
 QY 661 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGAAATACAGATTTTAAGCTTCATC 720  
 DB 718 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGAAATACAGATTTTAAGCTTCATC 777  
 QY 721 TCTTTGGCCCACTCCCTCCCTCAGATTGGCAAAATGAGTGGAGGAAATTTGGCTCA 780  
 DB 778 TCTTTGGCCCACTCCCTCCCTCAGATTGGCAAAATGAGTGGAGGAAATTTGGCTCA 837  
 QY 781 GAAGAAACATCTATATACATTTGAGAAAGAAAGTATGAGAGAGCCCAATGATAT 840  
 DB 838 GAAGAAACATCTATATACATTTGAGAAAGAAAGTATGAGAGAGCCCAATGATAT 897  
 QY 841 TATTTGATATGATGAGAGGAGGAGCAACCTTCAACCTTTGGCTTTGCTTTGCAATG 900  
 DB 898 TATTTGATATGATGAGAGGAGGAGCAACCTTCAACCTTTGGCTTTGCTTTGCAATG 957  
 QY 901 CCA 903  
 DB 958 CCA 960

RESULT 8  
 AAS22714 standard; cDNA; 1756 BP.

AAS22714;  
 24-OCT-2001 (first entry)

Human cDNA encoding a novel human protein #280.

Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;  
 immunomodulatory; cytoskeletal; neuroprotective; vulnary; nocotropic;  
 anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;  
 thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 Parkinson's disease; inflammatory disorder; cancer; aschma; osteoporosis;  
 tissue regeneration; immune disorder.

Homo sapiens.

WO200155437-A2.

02-AUG-2001.

25-JAN-2001; 2001WO-US02623.

25-JAN-2000; 2000US-0491404.

(HISE-) HISEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI: 2001-451939/48.

P-PSDB; AAU14409.

Isolated polypeptides useful for treating anti-inflammatory diseases,  
 nervous system disorders, and for regenerating bone and cartilage -

Claim 1; Page 693-694; 894pp: English.

The invention relates to polynucleotides encoding novel human  
 proteins or their active domains. The polypeptides, polynucleotides and  
 antibodies raised against the polypeptides are used in a method of  
 treatment of a mammal and prevention of disorders caused by the aberrant

CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/ elicit an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemic,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence encodes a protein of the invention.

Sequence 1756 BP; 441 A; 430 C; 413 G; 467 T; 5 other;

Query Match 100.0%; Score 903; DB 22; Length 1756;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-276;  
 Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 60  
 DB 43 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 102  
 QY 61 TCTTCAAGAGTGAATACAGAGCGAGTGGTCAATGCTTATCTGCTTCTTAC 120  
 DB 103 TCTTCAAGAGTGAATACAGAGCGAGTGGTCAATGCTTATCTGCTTCTTAC 162  
 QY 121 ACCCGAGCGCCCGGAGGAACTGGTCCGCTGCTGGGAGCAAGAGGAGCTGCTG 180  
 DB 163 ACCCGAGCGCCCGGAGGAACTGGTCCGCTGCTGGGAGCAAGAGGAGCTGCTG 222  
 QY 181 TTTGATGATGAGCAAGTGGTCTCAGAGCTGATGAAAGGATGTAATTTATGACATCC 240  
 DB 223 TTTGATGATGAGCAAGTGGTCTCAGAGCTGATGAAAGGATGTAATTTATGACATCC 282  
 QY 241 AGATCTGGCTTAATGAGGATTTCCGCAAGAGATGTCCTCTGACATAGAAATGTC 300  
 DB 283 AGATCTGGCTTAATGAGGATTTCCGCAAGAGATGTCCTCTGACATAGAAATGTC 342  
 QY 301 ACTTACGAGAGAGTGGGATCTACTGCTGCGGATCCAAATCCGAGGCTAATGATAT 360  
 DB 343 ACTTACGAGAGAGTGGGATCTACTGCTGCGGATCCAAATCCGAGGCTAATGATAT 402  
 QY 361 GAAAATTTAATCTGAGTTGGTTCATCAACCCAGCAAGTGCACCCCTGCACGACTCTG 420  
 DB 403 GAAAATTTAATCTGAGTTGGTTCATCAACCCAGCAAGTGCACCCCTGCACGACTCTG 462  
 QY 421 CAGAGAGCTTACATGACGACCTTTCCAGAGATGCTTACCAAGAGGAGACATGCCCCA 480  
 DB 463 CAGAGAGCTTACATGACGACCTTTCCAGAGATGCTTACCAAGAGGAGACATGCCCCA 522  
 QY 481 GAGACAGACACTGAGGAGCTCCCTGATATTAATTAACAACAATATCCACATTTGGC 540  
 DB 523 GAGACAGACACTGAGGAGCTCCCTGATATTAATTAACAACAATATCCACATTTGGC 582  
 QY 541 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 600  
 DB 583 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 642  
 QY 601 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 660  
 DB 643 AATGATTAACGGGACTCTAGATTGGCCATGACTTACGGGAGCTGGAGCAACCATCAGA 702  
 QY 661 GCTTTAATTTTCAATGATGATTTCTCATAGCAAGAGAAATACAGATTTTAAGCTTCATC 720







XX	Human 200 gene, preferentially expressed in TH1 cells.
DE	
XX	
KW	Human; 200 gene; T helper; T lymphocyte; T cell; TH1;
KW	receptor; differential expression; immune disorder; psoriasis;
KW	multiple sclerosis; insulin-dependent diabetes; antidiabetic;
KM	antipsoriatic; diagnosis; therapy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	42..947
FT	/+tag= a
FT	sig_peptide
FT	42..101
FT	/+tag= b
FT	mat_peptide
FT	102..944
FT	/+tag= c
XX	
PN	US6288218-B1.
PD	
PD	11-SEP-2001.
XX	
PF	25-SEP-1997; 97US-0937399.
XX	
PR	01-MAR-1996; 96US-0609583.
PR	03-MAR-1995; 95US-0398633.
XX	
PR	07-JUN-1995; 95US-0487748.
XX	
PA	(LEVI/) LEVINSON D. A.
XX	
PI	Levinson DA;
XX	
DR	WPI; 2001-647189/74.
XX	
DR	P-PSDB; AAM50223.
XX	
PT	Detecting 200 gene expression in a sample, useful for treating and
PT	diagnosing immune disorders, especially T lymphocyte-related disorders,
XX	
PT	comprises detecting the presence of a 200 gene product or an RNA
XX	
PT	encoding the 200 gene product
XX	
PS	Claim 1 (m); Fig 24A-D; 108pp; English.
XX	
CC	The present sequence is that of cDNA including the complete open
CC	reading frame of the human 200 gene. The encoded protein (see
CC	AAM50223) is a receptor of the Ig superfamily class. The cDNA was
CC	isolated from a human lymphocyte cDNA library using mouse 200 gene
CC	cDNA as probe. The nucleotide sequence contained within B. coli
CC	clone fent 200C is deposited as ATCC 69967. The human 200 gene is
CC	preferentially expressed in mature, fully differentiated T helper
CC	subpopulation TH1 cells relative to subpopulation TH2 cells. The
CC	gene can be used diagnostically or as a target for therapeutic
CC	intervention for the treatment of immune disorders. A claimed
CC	method for diagnosing a TH cell subpopulation-related immune
CC	disorder involves detecting the level of a 200 gene product, or
CC	an RNA encoding it, so that if the level differs from that in a
CC	control sample, the disorder is diagnosed. The disorder is
CC	especially a TH1 cell subpopulation-related immune disorder, such
CC	as multiple sclerosis, psoriasis or insulin-dependent diabetes
CC	(claimed). In addition to the 200 gene, the invention provides
CC	other genes that are differentially expressed within and among
CC	TH cells and TH cell subpopulations and which can be used in
CC	methods for the diagnosis, prognosis, evaluation and treatment of
CC	TH cell subpopulation-related disorders, for the identification of
CC	subjects exhibiting a predisposition to such conditions, for
CC	monitoring patients undergoing clinical evaluation for the
CC	treatment of such disorders, and for monitoring the efficacy of
CC	compounds used in clinical trials. Other immune disorders that can
CC	be treated/diagnosed include Crohn's disease, reactive arthritis,
CC	lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,
CC	grave's disease, contact dermatitis, graft rejection, graft versus
CC	host disease, sarcoidosis, atopic conditions, asthma, allergy,
CC	allergic rhinitis, food allergy, eosinophilia, conjunctivitis,
CC	glomerular nephritis, helminthic infection (e.g. leishmaniasis),

CC	viral infection (e.g. HIV), and bacterial infection (e.g.
CC	tuberculosis and lepromatous leprosy).
XX	
SQ	Sequence 2236 BP; 586 A; 518 C; 525 G; 606 T; 1 other;
Query Match	100.0%; Score 903; DB 22; Length 2236;
Best Local Similarity	100.0%; Pred. No. 7,7e-276;
Matches 903; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ATGTTTACACTCTCCCTTGATCGTGTCCGTCGTGCCTACTACTTAACAAGG 60
DB	
OY	42 ATGTTTACACTCTCCCTTGATCGTGTCCGTCGTGCCTACTACTTAACAAGG 101
DB	
OY	61 TCCTCAGAAGTGAATAACAGACGGAGGTGCGTCAAGATGCTATTCTGCCCTTCTAC 120
DB	
OY	102 TCCTCAGAAGTGAATAACAGACGGAGGTGCGTCAAGATGCTATTCTGCCCTTCTAC 161
DB	
OY	121 ACCCCAGCCCCCCAGGAGAACCTGTGCCGTCTGTCTGGGCAAGAAGCCTGTCTGTG 180
DB	
OY	162 ACCCCAGCCCCCCAGGAGAACCTGTGCCGTCTGTCTGGGCAAGAAGCCTGTCTGTG 221
DB	
OY	181 TTGTAATGTGCACACGTGNGCTCAGACCTGATGAAGAAGATGTAATTATTTGACATCC 240
DB	
OY	222 TTGTAATGTGCACACGTGNGCTCAGACCTGATGAAGAAGATGTAATTATTTGACATCC 281
DB	
OY	241 AGATACCTGGCTAAATATGGGAGTTTTCCGAAAGAGATGTGTCCCTGACCATAGAGAAATGTG 300
DB	
OY	282 AGATACCTGGCTAAATATGGGAGTTTTCCGAAAGAGATGTGTCCCTGACCATAGAGAAATGTG 341
DB	
OY	301 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGATCCAAATCCAGCATTAATGAATGAT 360
DB	
OY	342 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGATCCAAATCCAGCATTAATGAATGAT 401
DB	
OY	361 GAAAATTTAACCTGAAATGTGTGATCAACACGCAAGATCACCCCCTGACCCGACTCTG 420
DB	
OY	402 GAAAATTTAACCTGAAATGTGTGATCAACACGCAAGATCACCCCCTGACCCGACTCTG 461
DB	
OY	421 CAGAGAGACTTCACTGACAGCTTTTCCAAGATGCTTACACACGAGGAGCATGAGCCACAGA 480
DB	
OY	462 CAGAGAGACTTCACTGACAGCTTTTCCAAGATGCTTACACACGAGGAGCATGAGCCACAGA 521
DB	
OY	481 GAGACACAGACACTGGGGAGCTCCCTTGATATTAATCTTAACACAAATATTCACATTGGCC 540
DB	
OY	522 GAGACACAGACACTGGGGAGCTCCCTTGATATTAATCTTAACACAAATATTCACATTGGCC 581
DB	
OY	541 AATGAGTTACGGGACTCTAATTTGGCCAAATGACTTAACGGGACTCTGGAGCAACATTCAGA 600
DB	
OY	582 AATGAGTTACGGGACTCTAATTTGGCCAAATGACTTAACGGGACTCTGGAGCAACATTCAGA 641
DB	
OY	601 ATAGGCACTCAATCGAGACAGGAGTCTGTCTGGGCTGGGCTCTGGCTCTTATCTTGGGC 660
DB	
OY	642 ATAGGCACTCAATCGAGACAGGAGTCTGTCTGGGCTGGGCTCTGGCTCTTATCTTGGGC 701
DB	
OY	661 GCTTAAATTTTCAATGGTATTTCTATAGCAAGAAGATACAGAAATTTAAGCCTCATC 720
DB	
OY	702 GCTTAAATTTTCAATGGTATTTCTATAGCAAGAAGATACAGAAATTTAAGCCTCATC 761
DB	
OY	721 TCTTTGGCAACCTTCCTCTCAGAGATTGGCAAAATGAGTAGCAGAGGAAATTTGCTCA 780
DB	
OY	762 TCTTTGGCAACCTTCCTCTCAGAGATTGGCAAAATGAGTAGCAGAGGAAATTTGCTCA 821
DB	
OY	781 GAAGAAAAACATCTATACCATTTGAAGAGAAAGTATATGAAGTGAAGAGAGCCCAATGAGTAT 840
DB	
OY	822 GAAGAAAAACATCTATACCATTTGAAGAGAAAGTATATGAAGTGAAGAGAGCCCAATGAGTAT 881
DB	
OY	841 TATTGCTATGTCAAGACAGGACAGCAACCTTCACACACTTTGGTGTGTGCTTTGCAATG 900
DB	
OY	882 TATTGCTATGTCAAGACAGGACAGCAACCTTCACACACTTTGGTGTGTGCTTTGCAATG 941
DB	
OY	901 CCA 903
DB	
OY	942 CCA 944











**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 16:48:30 ; Search time 55.4846 Seconds  
(without alignments)  
7183.411 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903

Sequence: 1 atgtttcacatcttcctt.....gttcgtcgttcgaatgccca 903

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	903	100.0	903	US-08-829-525-37	Sequence 37, Appl
2	903	100.0	903	US-08-609-583A-37	Sequence 37, Appl
3	903	100.0	903	US-08-937-339-37	Sequence 37, Appl
4	903	100.0	903	US-09-310-367-37	Sequence 37, Appl
5	903	100.0	903	US-09-033-337-37	Sequence 37, Appl
6	903	100.0	903	US-09-464-231-37	Sequence 37, Appl
7	903	100.0	2236	US-08-829-525-23	Sequence 23, Appl
8	903	100.0	2236	US-08-609-583A-23	Sequence 23, Appl
9	903	100.0	2236	US-08-937-339-23	Sequence 23, Appl
10	903	100.0	2236	US-09-310-367-23	Sequence 23, Appl
11	903	100.0	2236	US-09-033-337-23	Sequence 23, Appl
12	903	100.0	2236	US-09-464-231-23	Sequence 23, Appl
13	456	50.5	2710	US-08-487-748A-8	Sequence 8, Appl
14	456	50.5	2710	US-08-480-070C-8	Sequence 8, Appl
15	456	50.5	2710	US-08-829-525-8	Sequence 8, Appl
16	456	50.5	2710	US-08-609-583A-8	Sequence 8, Appl
17	456	50.5	2710	US-08-937-339-8	Sequence 8, Appl
18	456	50.5	2710	US-09-310-367-8	Sequence 8, Appl
19	456	50.5	2710	US-09-033-337-8	Sequence 8, Appl
20	456	50.5	2710	US-09-464-231-8	Sequence 8, Appl
21	442.6	49.0	843	US-08-829-525-22	Sequence 22, Appl
22	442.6	49.0	843	US-08-609-583A-22	Sequence 22, Appl
23	442.6	49.0	843	US-08-937-339-22	Sequence 22, Appl
24	442.6	49.0	843	US-09-310-367-22	Sequence 22, Appl
25	442.6	49.0	843	US-09-033-337-22	Sequence 22, Appl
26	442.6	49.0	843	US-09-464-231-22	Sequence 22, Appl
27	64.4	7.1	1380	US-09-620-312D-59	Sequence 59, Appl

28	56.4	6.2	2093	1	US-08-287-001A-1	Sequence 1, Appl
29	56.4	6.2	2093	3	PCT-US95-09941-1	Sequence 1, Appl
30	37	4.1	578	5	US-09-085-198B-25	Sequence 25, Appl
31	35.2	3.9	2091	1	US-08-765-081-4	Sequence 4, Appl
32	35.2	3.9	2091	3	US-09-098-082-4	Sequence 4, Appl
33	35.2	3.9	2091	5	PCT-US95-06994-4	Sequence 4, Appl
34	35.2	3.9	8041	1	US-08-765-081-1	Sequence 1, Appl
35	35.2	3.9	8041	3	US-09-098-082-1	Sequence 1, Appl
36	35.2	3.9	8041	5	PCT-US95-06994-1	Sequence 1, Appl
37	35.2	3.9	87563	4	US-09-453-702B-57	Sequence 57, Appl
38	34.4	3.8	3502	2	US-08-724-394A-16	Sequence 16, Appl
39	32.6	3.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
40	31.8	3.5	2175	4	US-08-914-999-3	Sequence 3, Appl
41	31.8	3.5	90050	3	US-09-245-041-5	Sequence 5, Appl
42	31.4	3.5	2310	1	US-08-461-690B-1	Sequence 1, Appl
43	31.4	3.5	2311	1	US-08-103-445-1	Sequence 1, Appl
44	31.4	3.5	3257	5	PCT-US91-09784-1	Sequence 1, Appl
45	31.2	3.5	418	3	US-09-141-000-6	Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-08-829-525-37  
Sequence 37, Application US/08829525  
Patent No. 6084083  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829, 525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609, 583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487, 748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398, 633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-829-525-37  
Query Match 100.0%; Score 903; DB 3; Length 903;



QY 421 CAGAGAGACTTCACTGACGCTTTTCCAGATGCTTACCACGAGGAGCATGCCCCACA 480  
DB 421 CAGAGAGACTTCACTGACGCTTTTCCAGATGCTTACCACGAGGAGCATGCCCCACA 480  
QY 481 GAGACACAGACACTGGGAGGCTTCCGTATATAATCTAACCAATAATTCACATTTGCC 540  
DB 481 GAGACACAGACACTGGGAGGCTTCCGTATATAATCTAACCAATAATTCACATTTGCC 540  
QY 541 AATAGATTAAGGAGCTCTAGATTGGCCAAATGACTTACGGGACTCTGGGCAACCTACGA 600  
DB 541 AATAGATTAAGGAGCTCTAGATTGGCCAAATGACTTACGGGACTCTGGGCAACCTACGA 600  
QY 601 ATAGGACATCTACATCGAGAGGAGATCTGTGCGGCTGGCTGGCTCTTATCTTCCGC 660  
DB 601 ATAGGACATCTACATCGAGAGGAGATCTGTGCGGCTGGCTGGCTCTTATCTTCCGC 660  
QY 661 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTTCATC 720  
DB 661 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTTCATC 720  
QY 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGCAATGACAGAGGAAATTCGCTCA 780  
DB 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGCAATGACAGAGGAAATTCGCTCA 780  
QY 781 GAAGAAAACATCTATATACATTGAAGAAACGTAATGAAAGTGAAGAGAGAGCCCAATGATAT 840  
DB 781 GAAGAAAACATCTATATACATTGAAGAAACGTAATGAAAGTGAAGAGAGAGCCCAATGATAT 840  
QY 841 TATTGCTATGTGACGAGGAGGAGCAACCTTCAACCTTTGGCTTGTCTGCTTTCATG 900  
DB 841 TATTGCTATGTGACGAGGAGGAGCAACCTTCAACCTTTGGCTTGTCTGCTTTCATG 900  
QY 901 CCA 903  
DB 901 CCA 903

RESULT 3  
US-08-937-399-37  
; Sequence 37, Application US/08937399  
; Patent No. 6288218  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,399  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 903 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; US-08-937-399-37

Query Match 100.0%; Score 903; DB 3; Length 903;  
Best Local Similarity 100.0%; Pred. No. 1,3e-297;  
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTTCACATCTTCCCTTGAAGTGTCTGCTGCTGCTGCTACTACTTACAAAG 60  
DB 1 ATGTTTTCACATCTTCCCTTGAAGTGTCTGCTGCTGCTGCTACTACTTACAAAG 60  
QY 61 TCTTCAGAGTGAATACAGAGCGAGTGGTCAAAATGCTTATCTGCTCTTCTAC 120  
DB 61 TCTTCAGAGTGAATACAGAGCGAGTGGTCAAAATGCTTATCTGCTCTTCTAC 120  
QY 121 ACCCCAGCGCGCCAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 121 ACCCCAGCGCGCCAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGAGATGTAATTTGACATCC 240  
DB 181 TTTGAATGTGGCAACGTGTGCTCAGAGACTGATGAAGAGATGTAATTTGACATCC 240  
QY 241 AGATACCTGGCTAAATGGGATTTCCGCAAGAGATGTCTCTGACCATAGAGATGTG 300  
DB 241 AGATACCTGGCTAAATGGGATTTCCGCAAGAGATGTCTCTGACCATAGAGATGTG 300  
QY 301 ACTCTAGAGACATGGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 ACTCTAGAGACATGGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 GAAAAATTTAATCTGAAGTGTGATCAAAACCAAGCAAGTCAACCTGCAACGACTGTG 420  
DB 361 GAAAAATTTAATCTGAAGTGTGATCAAAACCAAGCAAGTCAACCTGCAACGACTGTG 420  
QY 421 CAGAGAGACTTCACTGACGCTTTTCCAGATGCTTACCACGAGGAGCATGCCCCACA 480  
DB 421 CAGAGAGACTTCACTGACGCTTTTCCAGATGCTTACCACGAGGAGCATGCCCCACA 480  
QY 481 GAGACACAGACACTGGGAGGCTTCCGTATATAATCTAACCAATAATTCACATTTGCC 540  
DB 481 GAGACACAGACACTGGGAGGCTTCCGTATATAATCTAACCAATAATTCACATTTGCC 540  
QY 541 AATAGATTAAGGAGCTCTAGATTGGCCAAATGACTTACGGGACTCTGGGCAACCTACGA 600  
DB 541 AATAGATTAAGGAGCTCTAGATTGGCCAAATGACTTACGGGACTCTGGGCAACCTACGA 600  
QY 601 ATAGGACATCTACATCGAGAGGAGATCTGTGCGGCTGGCTGGCTCTTATCTTCCGC 660  
DB 601 ATAGGACATCTACATCGAGAGGAGATCTGTGCGGCTGGCTGGCTCTTATCTTCCGC 660  
QY 661 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTTCATC 720  
DB 661 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAAGATACAGAAATTTAAGCTTCATC 720  
QY 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGCAATGACAGAGGAAATTCGCTCA 780  
DB 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGCAATGACAGAGGAAATTCGCTCA 780  
QY 781 GAAGAAAACATCTATATACATTGAAGAAACGTAATGAAAGTGAAGAGAGAGCCCAATGATAT 840  
DB 781 GAAGAAAACATCTATATACATTGAAGAAACGTAATGAAAGTGAAGAGAGAGCCCAATGATAT 840











```

OY 601 ATAGCATCTACATCGACAGCAGGATCTGCTGGCTGGCTGGCTCTTATCTTCGCG 668
Db 642 ATAGCATCTACATCGACAGCAGGATCTGCTGGCTGGCTGGCTCTTATCTTCGCG 701
OY 661 GCTTAAATTTTCAAAATGATATTCTCATAGCAAGAGAAATACAGATTTTAAGCTCATC 720
Db 702 GCTTAAATTTTCAAAATGATATTCTCATAGCAAGAGAAATACAGATTTTAAGCTCATC 761
OY 721 TCTTTGGCCACCTCCCTCCCTCAGATTTGGCAAAATGCAATAGCAGAGGAAATTCGCTCA 780
Db 762 TCTTTGGCCACCTCCCTCCCTCAGATTTGGCAAAATGCAATAGCAGAGGAAATTCGCTCA 821
OY 761 GAAGAAAACATCATATATACATTTGAAGAGAACGTTATGATGAGAGAGCCCAATGAGTAT 840
Db 822 GAAGAAAACATCTATATCCATTTGAAGAGAACGTTATGATGAGAGAGCCCAATGAGTAT 881
OY 841 TATTGCTATGTCAGCAGCAGGAGCAACCTCTCAACACTTTGGGTGTTCGCTTTCAGATG 900
Db 882 TATTGCTATGTCAGCAGCAGGAGCAACCTCTCAACACTTTGGGTGTTCGCTTTCAGATG 941
OY 901 CCA 903
Db 942 CCA 944

RESULT 9
US-08-937-399-23
Sequence 23, Application US/08937399
Patent No. 6288218
GENERAL INFORMATION:
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OR INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,399
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/609,583
FILING DATE: 01-MAR-1996
APPLICATION NUMBER: US 08/487,748
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/398,633
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Cortuzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2236 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA

```

```

FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 42..944
; OTHER INFORMATION: Human 200 gene nucleotide
; OTHER INFORMATION: sequence
US-08-937-399-23

Query Match      100.0%; Score 903; DB 3; Length 2236;
Best Local Similarity 100.0%; Pred. No. 2,3e-297;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 ATGTTTTCACATCTTCCCTTTGACTGTGTCCTGCTGCTGCTGCTACTACTACTTAACAAG 60
DB 42 ATGTTTTCACATCTTCCCTTTGACTGTGTCCTGCTGCTGCTGCTGCTACTACTACTTAACAAG 101
OY 61 TTCCTCAGAAAGGGAATACAGAGCGGAAGGTGGGTGCAATATGCTATCTAGCCCTGCTTAC 120
DB 102 TCCTCAGAAAGGGAATACAGAGCGGAAGGTGGGTGCAATATGCTATCTAGCCCTGCTTAC 161
OY 121 ACCCCAGCCGCCCCAGGGGAACCTGTGCGCGCTCTGCTGCGGCAAGAGACCTGTCTG 180
DB 162 ACCCCAGCCGCCCCAGGGGAACCTGTGCGCGCTCTGCTGCGGCAAGAGACCTGTCTG 221
OY 181 TTTGAAATGTGGCAACGTGTGTCTCAGACTGATGAAAGGATGTGTAATTTATTTGACATCC 240
DB 222 TTTGAAATGTGGCAACGTGTGTCTCAGACTGATGAAAGGATGTGTAATTTATTTGACATCC 281
OY 241 AGATACGTGGCTAAATGGGGATTTCCGCAAGAGATGTGTCCCTGACCATAGAGATGTG 300
DB 282 AGATACGTGGCTAAATGGGGATTTCCGCAAGAGATGTGTCCCTGACCATAGAGATGTG 341
OY 301 ACTCTAGCAGACAGTGGGATCTACTGCTGCGGATCCAAATCCAGGCATTAATGAATGAT 360
DB 342 ACTCTAGCAGACAGTGGGATCTACTGCTGCGGATCCAAATCCAGGCATTAATGAATGAT 401
OY 361 GAAAAAATTTAACTGGAAGTTGGTATCAAAACAGCAAGAGTCAACCCCTGCACCGACTCG 420
DB 402 GAAAAAATTTAACTGGAAGTTGGTATCAAAACAGCAAGAGTCAACCCCTGCACCGACTCG 461
OY 421 CAGAGAGACTTCACTGCAGCCCTTTCCAGAGATGCTTACCAACGGGGACATGGCCACGA 480
DB 462 CAGAGAGACTTCACTGCAGCCCTTTCCAGAGATGCTTACCAACGGGGACATGGCCACGA 521
OY 481 GAGACACAGACACTGGGGAGGCTCCCTGATATTAATCTTAACAACAATATCCACTTGGCC 540
DB 522 GAGACACAGACACTGGGGAGGCTCCCTGATATTAATCTTAACAACAATATCCACTTGGCC 581
OY 541 AATGAGTTACGGGACTCTAGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACATACGA 600
DB 582 AATGAGTTACGGGACTCTAGATTTGGCCAAATGACTTAACGGGACTCTGAGCAACATACGA 641
OY 601 ATAGGCATCTACATCGAGACAGAGGATCTGTGTCTGGGCTGCGCTCTTATCTTCCGC 660
DB 642 ATAGGCATCTACATCGAGACAGAGGATCTGTGTCTGGGCTGCGCTCTTATCTTCCGC 701
OY 661 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAGAGATACAGAAATTTAAGCCATAC 720
DB 702 GCTTTAATTTTCAATGGTATTTCTCATAGCAAGAGAGATACAGAAATTTAAGCCATAC 761
OY 721 TCTTTGGCCAACTTCCCTCTCCCTCAGAGATTGGCAATGCAATGACAGAGGAATTGCTCA 780
DB 762 TCTTTGGCCAACTTCCCTCTCCCTCAGAGATTGGCAATGCAATGACAGAGGAATTGCTCA 821
OY 781 GAAAGAAAACATCTATACCATTTGAAGAGAAACGATATGAAAGTGGAGAGGCCAATAGTAT 840
DB 822 GAAAGAAAACATCTATACCATTTGAAGAGAAACGATATGAAAGTGGAGAGGCCAATAGTAT 881
OY 841 TATTGCTATGTCAAGCAGCAGCAGCAACCTCTACACACTTTGGTGTGTGCTTTGCATG 900
DB 882 TATTGCTATGTCAAGCAGCAGCAGCAACCTCTACACACTTTGGTGTGTGCTTTGCATG 941
OY 901 CCA 903
|||

```









QY	240	---CAGTACTGGCTAATGGGGAATTTCCGCAAGGAGAATGTGCCCTGACCATAGAAAT	297
Db	280	ACCAAGTATACAGGTAAAGGCGCATCTCAACAAAGGAGATGTCTCTGATCATTAAGAT	339
QY	298	GTGACTCTAGCAGACAGTGGGATCTTACTGTGCGCGGATCCAAATCCGAGGCATTAAT	357
Db	340	GTGACTCTGGATGACCATGGGACCTTACTGTGCGAGGATACAGTGTCCCTGCTTATGAT	399
QY	358	GATGAAAAATTTAACTGAAGTTGGATCATAAACCGCCAAAGTACCCCTGCACCGACT	417
Db	400	GATAAAAATTTAGAACTGAAATTTAGCATTAAGCAGCCAAAGTCACTTCAGCTCAGACT	459
QY	418	CTGCAGAGAGACTTCACTGCAAGCTTTCCAGAGATGCTTACCAACAGGGGACATGGCCA	477
Db	460	GCCCACTGGGGACTCTACTACAGCTTCTCCAAAGACCTTACCAACAGGAGAAATGG--T	516
QY	478	GGAGAGACACAGACACTGGGGAGCTTCCCTGATTTAATTTAAACAATAATCCACTTG	537
Db	517	TGAGAGACACAGACACTGGGTGAGCCCTCCATTAATAACAAATGGAAACAATAATTTCCACTGG	576
QY	538	GCCAAATGAATTACGGGACTCTAGATTGGCCAAATGACTTACGGGACTCTGGAGCAACATC	597
Db	577	GCTGATGAATTA-----GGACTCTGGAGAAAGCATC	609
QY	598	AGAAATAGGCATCTACATCGAGACGAGATCTGTGGCTGGCTCTGCTTTATCTTC	657
Db	610	AAACACTGCTATCCACATTTGGAGGGAGTCTGCTGGGTGGACCTGGGACATTAATCATT	669
QY	658	GGCGCTTTAATTTCAATGTGTATTCATCTAGCAAGAGAAAGATACGAATTTAAGCTC	717
Db	670	GGTGTCTTATCTTTAAATGTGATTTCTCTTAAGAAAAGAAATTTATCGAATTTGAGCTT	729
QY	718	ATCTCTTTGGCCAACCTCTCTCCCTCAGATTGGCAATTCAGTAGCAGAGGGAATTCGC	777
Db	730	ATTAACTGGCCAACCTGCTCCAGGAGGGGTGGCAATTCAGAGACAGTCAGGAATTCGC	789
QY	778	TCAGAAAGAAACATCATATCAATTGAAGAAAGCTATATCAATTTGAGAGGCCAATGAG	837
Db	790	TCTGAGAAATATCTACACCATCGAGAGAAAGTATATGAATGAGAAATTTAAATGAG	849
QY	838	TATTAATGCTATGTCAAGCAGACGAGCAACCTTCAACCTTTGGGTGTGCTTTGCA	897
Db	850	TACTACTGTACTGATCAAGCAGCGACGACCATCTGACCGGCTCTGGAGCTGCCATTTTAA	909
QY	898	ATGC 901	
Db	910	AGGC 913	

RESULT 14  
 US-08-480-070C-8  
 Sequence 8, Application US/08480070C  
 Patent No. 6066698  
 GENERAL INFORMATION:  
 APPLICANT: Levinson, Douglas A.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 NUMBER OF SEQUENCES: 24  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,070C

```

? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A.
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 7853-024
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2710 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 40..885
?
US-08-480-070C-8

```

[illegible]





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 12:17:01 ; Search time 1977.2 Seconds  
(without alignments)  
11099.996 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903

Sequence: 1 atgttcacatctccctt.....gtgtcgttcgaatgcca 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estp1:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	901	99.8	1201	13	BX417090 BX417090
2	606	67.1	735	14	CA307836 CA307836
3	583	64.6	583	13	BU789118 BU789118
4	568.2	62.9	863	14	CD520418 CD520418

5	525	58.1	554	13	BQ581947
6	518.4	57.4	575	9	AU279844
7	499.4	55.3	746	14	CD522575
8	477.6	52.9	869	14	CD532271
9	451.4	50.0	796	12	BI906710
10	421	46.6	518	14	CB155731
11	397.4	44.0	572	10	AG709094
12	396.6	43.9	837	10	BG545730
13	349.8	38.7	519	12	BI834671
14	345	38.2	1370	13	BQ278485
15	304.8	33.8	633	14	BY741594
16	277.8	30.8	677	14	BY748679
17	199.8	22.1	535	12	BI541253
18	191	21.2	606	9	AV664631
19	150.6	16.7	429	9	AV664630
20	148	16.4	655	13	BQ581620
21	147.6	16.3	557	12	BM255779
22	137.2	15.2	350	13	BY185793
23	118.8	13.2	296	10	BB572107
24	111.6	12.4	634	13	BU238988
25	111.6	12.4	694	9	AJ452225
26	111.6	12.4	776	9	AJ452245
27	111.6	12.4	807	9	AJ445463
28	102	11.3	754	13	BU320531
29	100.4	11.1	900	13	BU346895
30	100.2	11.1	544	13	BX264550
31	97	10.7	414	13	BQ484385
32	91	10.1	820	13	BU207995
33	82.4	9.1	615	13	BX265767
34	75	8.3	641	13	BQ400478
35	75	8.3	828	14	CD254919
36	74	8.2	568	12	BM261288
37	74	8.2	901	13	BU148837
38	73.4	8.1	525	12	BI443222
39	72.6	8.0	722	9	AM604297
40	72.4	8.0	754	14	CB938882
41	71.8	8.0	688	13	BU206552
42	71	7.9	457	10	BE722324
43	70.2	7.8	761	10	BG168159
44	67.6	7.5	569	9	AJ454449
45	67.6	7.5	751	13	BX114328

## ALIGNMENTS

RESULT 1  
BX417090  
LOCUS  
DEFINITION BX417090 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB0041C05  
5-PRIME, mRNA sequence.  
ACCESSION BX417090  
VERSION BX417090.1 GI:30765665  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNALS Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2998.r for more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgl-bin/cluster.cgi?seq=CS0DB004AB03QPIcluster=2998.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600



Best Local Similarity 89.2%; Pred. No. 3,1e-172;  
Matches 700; Conservative 0; Mismatches 1; Indels 84; Gaps 1;

119 ACACCCGAGCGCCGCCGAGAACTCGTCCGCTGCTGGGGGCAAGAGCCCTCTCTG 178  
Db AACCCCGAGCGCCGCCGAGAACTCGTCCGCTGCTGGGGGCAAGAGCCCTCTCTG 178  
735 AACCCCGAGCGCCGCCGAGAACTCGTCCGCTGCTGGGGGCAAGAGCCCTCTCTG 676  
179 TGTTCATGTCGCAACGTGCTGCTCAGAGCTGATGAAGAGGATGATGATGATGAT 238  
Db TGTTCATGTCGCAACGTGCTGCTCAGAGCTGATGAAGAGGATGATGATGATGAT 616  
675 TGTTCATGTCGCAACGTGCTGCTCAGAGCTGATGAAGAGGATGATGATGATGAT 616  
239 CCAAGTACTGCTTAATGAGGATTTCCGCAAGAGAGATGCTGCTGATGATGAGATG 298  
Db CCAAGTACTGCTTAATGAGGATTTCCGCAAGAGAGATGCTGCTGATGATGAGATG 556  
615 CCAAGTACTGCTTAATGAGGATTTCCGCAAGAGAGATGCTGCTGATGATGAGATG 556  
299 TGACTCTGAGCAGAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358  
Db TGACTCTGAGCAGAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496  
555 TGACTCTGAGCAGAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496  
359 ATGAAATTTTAACTGAGTGTGCTCATCAACAGCAAGGTACCCCTGACCGACTG 418  
Db ATGAAATTTTAACTGAGTGTGCTCATCAACAGCAAGGTACCCCTGACCGACTG 464  
495 ATGAAATTTTAACTGAGTGTGCTCATCAACAGCAAGGTACCCCTGACCGACTG 464  
419 TGACAGAGACTTCACTGCAAGCTTTCCAGAGATGCTTACACAGGGGACATGCGC 478  
Db TGACAGAGACTTCACTGCAAGCTTTCCAGAGATGCTTACACAGGGGACATGCGC 463  
463 -----CCAG 460  
479 CAGAGACACAGACATGCGGGAGCTCCTCGATATATATTAACACAAATATCCATG 538  
Db CAGAGACACAGACATGCGGGAGCTCCTCGATATATATTAACACAAATATCCATG 459  
459 CAGAGACACAGACATGCGGGAGCTCCTCGATATATATTAACACAAATATCCATG 400  
539 CCAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 598  
Db CCAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 399  
399 CCAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 340  
599 GAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 658  
Db GAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 339  
339 GAATGAGTTCAGGAGCTCTAGATGCGCAATGATTAACGAGGACTGAGAGCAAC 280  
659 GCGCTTAAATTTTCAATGATGATTTCTCATAGCAAGAGATTAAGATTAAGCTCA 718  
Db GCGCTTAAATTTTCAATGATGATTTCTCATAGCAAGAGATTAAGATTAAGCTCA 279  
279 GCGCTTAAATTTTCAATGATGATTTCTCATAGCAAGAGATTAAGATTAAGCTCA 220  
719 TCTCTTGGCCCAACCTCCTCCCTCAGAGTGGCAATGCAATGAGAGGATTTGCT 778  
Db TCTCTTGGCCCAACCTCCTCCCTCAGAGTGGCAATGCAATGAGAGGATTTGCT 219  
219 TCTCTTGGCCCAACCTCCTCCCTCAGAGTGGCAATGCAATGAGAGGATTTGCT 160  
779 CAGAGAAACATCTATACATGATGAAGAGACATGATGATGATGATGATGATGAT 838  
Db CAGAGAAACATCTATACATGATGAAGAGACATGATGATGATGATGATGATGAT 159  
159 CAGAGAAACATCTATACATGATGAAGAGACATGATGATGATGATGATGATGAT 100  
839 ATTATGCTATGTCAGCAGAGGAGCAACCTCAACACCTTTGGGTTGCTTGCAT 898  
Db ATTATGCTATGTCAGCAGAGGAGCAACCTCAACACCTTTGGGTTGCTTGCAT 99  
99 ATTATGCTATGTCAGCAGAGGAGCAACCTCAACACCTTTGGGTTGCTTGCAT 40  
899 TGCCA 903  
Db TGCCA 35

RESULT 3  
BU789118 583 bp mRNA linear EST 11-OCT-2002  
LOCUS 104002.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128595  
DEFINITION 3', similar to TR:054947 054947 KIDNEY INJURY MOLECULE-1 PRECURSOR  
BU789118  
ACCESSION BU789118.1 GI:23838264  
VERSION BU789118  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 583)  
AUTHORS Melton,D., Brown,J., Keny,G., Permutt,A., Lee,C., Kaestner,K.,  
Lemishka,I., Scaer,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Biscain,A.,  
Schmitt,A., Theising,B., Ritter,E., Ronko,L., Bennett,J., Gardenas  
,M., Gibbons,M., McCann,R., Cole,R., Tsagarelashvili,R., Williams,T.,  
Jackson,Y. and Bowers,Y.  
TITLE Endocrine Pancreas Consortium  
JOURNAL Unpublished  
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biolup.harvard.edu  
Library was constructed by Dr. J. Ferrer in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Possible reversed clone: similarity on wrong strand  
Seq primer: -400P from Gibco  
High quality sequence stop: 460.  
Location/Qualifiers  
1..583  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6128595"  
/tissue\_type="insulinoma"  
/lab\_host="pDH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/note="Organ: pancreas; Vector: pBluescript SK-. Site.1:  
XhoI; Site.2: EcoRI; Constructed with lambda ZAPI system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further  
information on this library (Metabolism Division, Permutt  
Laboratory, Washington University School of Medicine, Box  
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
is a Washington University Pancreas EST project library."

BASE COUNT 152 a 152 c 148 g 131 t  
ORIGIN

Query Match 64.6%; Score 583; DB 13; Length 583;  
Best Local Similarity 100.0%; Pred. No. 2.6e-165;  
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 TCACATCTTCCCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 66  
Db TCACATCTTCCCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60  
67 GAATGGAATACAGAGCGAGAGTGTGCTGAGATGCTTATGCTGCTGCTTCAACCCCA 126  
Db GAATGGAATACAGAGCGAGAGTGTGCTGAGATGCTTATGCTGCTGCTTCAACCCCA 61  
61 GAATGGAATACAGAGCGAGAGTGTGCTGAGATGCTTATGCTGCTGCTTCAACCCCA 120  
127 GCGGCCCGAGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
Db GCGGCCCGAGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121  
121 GCGGCCCGAGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
167 TGTGGCAACGTGTGCTCAGAGCTGATGAAGAGATGATGATGATGATGATGATGAT 246  
Db TGTGGCAACGTGTGCTCAGAGCTGATGAAGAGATGATGATGATGATGATGATGAT 181  
181 TGTGGCAACGTGTGCTCAGAGCTGATGAAGAGATGATGATGATGATGATGATGAT 240  
247 TGCTTAATGAGGATTTCCGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306  
Db TGCTTAATGAGGATTTCCGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
241 TGCTTAATGAGGATTTCCGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
307 GCAGACAGTGGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366







DB 356 ACTTATGACAGACGTGGATCTACTGCTCCGGATCCAAATCCAGACATAATGAAT 415  
QY 361 GAAAAATTTAACTGTAAGTTGTGATCAAAACGACCAAGGTCAACCCCTGACCGACTCTG 420  
DB 416 GAAAAATTTAACTGTAAGTTGTGATCAAAACGACCAAGGTCAACCCCTGACCGACTCTG 475  
QY 421 CAGAGAGACTTCACTGACGCTTTCCAGAGATGCTTACCAACGAGGACATGAGCCACGCA 480  
DB 476 CAGAGAGACTTCACTGACGCTTTCCAGAGATGCTTACCAACGAGGACATGAGCCACGCA 535  
QY 481 GAGACACAGACACTGGGAGGAGCTCCCTGATATTAATCTAA 520  
DB 536 GAGACACAGACACTGGGAGGAGCTCCCTGATATTAATCTAA 575  
RESULT 7  
CD522575 746 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_14353511 NIH\_MGC\_191 Homo sapiens cDNA clone  
DEFINITION IMAGE:30411065 5', mRNA sequence.  
ACCESSION CD522575  
VERSION CD522575.1 GI:31454293  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 746)  
TITLE NIH-MGC http://mgs.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM202 row: d column: 18  
High quality sequence stop: 592.  
Location/Qualifiers  
1..746  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30411065"  
/issue\_type="Pooled"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_id="NIH\_MGC\_191"  
/note="Vector: pDNR-LIB, site 1: SfiI (ggccatctggcc);  
site 2: SfiI (ggcgcctcgcc); Library is oligo-dT primed  
and directionally cloned. PBMC - Peripheral Blood  
Mononuclear Cells. RNA was pooled from 3/6hour stimulation  
with PMA adn Ionomycin. 5' and 3' adaptors were used in  
cloning as follows: 5' adaptor sequence:  
5'-CACGCGCATTTAGCC-3' and 3' adaptor sequence:  
5'-ATCTTAGAGCCGAGGCGCGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.69  
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH MGC Library."

BASE COUNT 191 a 185 c 192 g 177 t 1 others  
ORIGIN  
Query Match 55.3%; Score 499.4; DB 14; Length 746;  
Best Local Similarity 97.6%; Pred. No. 7e-140;

Matches 518; Conservative 0; Mismatches 11; Indels 2; Gaps 1;  
QY 1 ATGTTTCAACATCTTCCCTTTGACTGTGTCTCTGCTCTGCTACTTACTTACAAAG 60  
DB 143 ATGTTTCAACATCTTCCCTTTGACTGTGTCTCTGCTCTGCTACTTACTTACAAAG 202  
QY 61 TCCTCAGAGTGAATATACAGAGCGGAGGTGGTCAAGATGCTATGCTGCTCTCAG 120  
DB 203 TCCTCAGAGTGAATATACAGAGCGGAGGTGGTCAAGATGCTATGCTGCTCTCAG 262  
QY 121 ACCCCAGCCGCCCCAGAGGAACCTGTGCTGCTGCTGCTGAGGAGAAAGAGCTGTCTGTG 180  
DB 263 ACCCCAGCCGCCCCAGAGGAACCTGTGCTGCTGCTGCTGAGGAGAAAGAGCTGTCTGTG 322  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGAGATGATGAAGGAGATGTAATTAATGACATCC 240  
DB 323 TTTGAATGTGGCAACGTGTGCTCAGAGATGATGAAGGAGATGTAATTAATGACATCC 382  
QY 241 AGATTAAGTGAATATGAGGAGATTTCCGCAAGAGAGATGTCCTGACATAGAGATGTG 300  
DB 383 AGATTAAGTGAATATGAGGAGATTTCCGCAAGAGAGATGTCCTGACATAGAGATGTG 442  
QY 301 ACTCTAGCAGACAGTGGGATCTACTGCTGCTCCGATCCAAATCCAGAGCAATTAATGAAT 360  
DB 443 ACTCTAGCAGACAGTGGGATCTACTGCTGCTCCGATCCAAATCCAGAGCAATTAATGAAT 502  
QY 361 GAAAAATTTAACTGTAAGTTGTGATCAAAACGACCAAGGTCAACCCCTGACCGACTCTG 420  
DB 503 GAAAAATTTAACTGTAAGTTGTGATCAAAACGACCAAGGTCAACCCCTGACCGACTCTG 562  
QY 421 CAGAGAGACTTCACTGACGCTTTCCAGAGATGCTTACCAACGAGGAGATGAGCCACGCA 480  
DB 563 CAGAGAGACTTCACTGACGCTTTCCAGAGATGCTTACCAACGAGGAGATGAGCCACGCA 622  
DB 481 GAGACACAGACACTGGGAGGAGCTCCCTGATATTAATCTAAACAAATAT 529  
DB 623 GAGACACAGACACTGGGAGGAGCTCCCTGATATTAATCTAAACAAATAT 673  
RESULT 8  
CD523271 869 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_14356479 NIH\_MGC\_191 Homo sapiens cDNA clone  
DEFINITION IMAGE:30410457 5', mRNA sequence.  
ACCESSION CD523271  
VERSION CD523271.1 GI:31454989  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 869)  
TITLE NIH-MGC http://mgs.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM200 row: k column: 10  
High quality sequence stop: 435.  
Location/Qualifiers  
1..869  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source  
1..869  
/organism="Homo sapiens"  
/mol\_type="mRNA"





QY 528 ATCCA 532  
Db 791 TTCCA 795

RESULT 10  
CB155731 518 bp mRNA linear EST 29-JAN-2003  
LOCUS K-EST0214151 B2N807043 Homo sapiens cDNA clone B2N807043-7-D07 5',  
DEFINITION mRNA sequence.  
ACCESSION CB155731  
VERSION CB155731.1 GI:28140844  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 518)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,D.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongjung@mail.kribb.re.kr  
Plate: 7 row: D column: 07  
High quality sequence stop: 518.  
Location/Qualifiers  
1..518  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="B2N807043-7-D07"  
/sex="M"  
/lab\_host="Top10F"  
/clone\_11b="B2N807043"  
/note="Organ: Brain; Vector: pCNS-D2; Site 1: EcoRI;  
Site\_2: NotI. The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

BASE COUNT 128 a 128 c 132 g 130 t

ORIGIN

Query Match 46.6%; Score 421; DB 14; Length 518;  
Best Local Similarity 100.0%; Pred. No. 3,4e-116;  
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTTCACATCTCCCTTGAAGTGTCTGCTGCTGCTGCTACTACTTACAGG 60  
|||  
Db 98 ATGTTTTCACATCTCCCTTGAAGTGTCTGCTGCTGCTGCTACTACTTACAGG 157  
|||

QY 61 TCCTCAGAGTGAATACAGAGCGAGTGGTCAGAGTCTATCTGCTGCTTCTTAC 120  
|||  
Db 158 TCCTCAGAGTGAATACAGAGCGAGTGGTCAGAGTCTATCTGCTGCTTCTTAC 217  
|||

QY 121 ACCCCAGCCGCCAGGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
|||

Db 218 ACCCAGCCGCCAGGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277  
|||

QY 181 TTGAATGTGGCAACGTGTGCTCAGACTGATGAAGGATGTGATTTATGACATCC 240  
|||

Db 278 TTGAATGTGGCAACGTGTGCTCAGACTGATGAAGGATGTGATTTATGACATCC 337  
|||

QY 241 AGATCTGCTTAATATGGGATTTCCGCAAGAGATGTGCTCCCTGACCATAGAGATG 300  
|||

Db 338 AGATCTGCTTAATATGGGATTTCCGCAAGAGATGTGCTCCCTGACCATAGAGATG 397  
|||

QY 301 ACTCTAGCAGACAGTGGGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
|||

Db 398 ACTCTAGCAGACAGTGGGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457  
|||

QY 361 GAAAAATTATTAACCTGAAGTTGTGATCAACAGCAGCCAGGTACCCCTGACCGACTCG 420  
|||

Db 458 GAAAAATTATTAACCTGAAGTTGTGATCAACAGCAGCCAGGTACCCCTGACCGACTCG 517  
|||

QY 421 C 421  
|

Db 518 C 518

RESULT 11  
AL709094 572 bp mRNA linear EST 22-MAR-2002  
LOCUS DKFZp686F1557.1 686 (synonym: h1cc3) Homo sapiens cDNA clone  
DEFINITION DKFZp686F1557.5, mRNA sequence.  
ACCESSION AL709094  
VERSION AL709094.1 GI:19692449  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 572)  
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,W., Mewes,H.W., Weil,B. and  
Wiemann,S.  
TITLE EST (Ottenwaelder,B., Obermaier,B., Mewes,H.W., Weil,B. and Wiemann,  
S.)  
JOURNAL Unpublished  
COMMENT Contact: Ottenwaelder B  
MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Medigenomix (Martinsried/Germany) within the cDNA  
sequencing consortium of the German Genome Project. No sl sequence  
available.  
This clone (DKFZp686F1557) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..572  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686F1557"  
/cissue\_type="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_11b="686 (synonym: h1cc3)"  
/note="Vector: pTriplEx2; Site\_1: SfiI; Site\_2: SfiIb;  
cDNA-collection"

BASE COUNT 141 a 137 c 152 g 142 t

ORIGIN

Query Match 44.0%; Score 397.4; DB 9; Length 572;  
Best Local Similarity 99.7%; Pred. No. 5,2e-109;  
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGTTTACAACTTCCCTTGACCTGTGCCGCTGCTGCTGCTACTTACAAAG	60
Db	174	ATGTTTACAACTTCCCTTGACCTGTGTCTGCTGCTGCTGCTACTTACAAAG	233
Qy	61	TCCTCAGAACTGGAATACAGACGCGAGGTGCGTCAGAAATGCTTATCGCTTCTTAC	120
Db	234	TCCTCAGAAAGTGAATATACAGACGCGAGGTGCGTCAGAAATGCTTATCGCTTCTTAC	293
Qy	121	ACCCCAAGCGGCCCCCAGGGAACTCGTGCCGCTGTGCTGGGGCAAAAGAGCTGTCTGTG	180
Db	294	ACCCCAAGCGGCCCCCAGGGAACTCGTGCCGCTGTGCTGGGGCAAAAGAGCTGTCTGTG	353
Qy	181	TTTGAATGTGGCAAGCTGTGTGCTCAGACCTGAATGAAGGAAATGAATTAATTGACATCC	240
Db	354	TTTGAATGTGGCAAGCTGTGTGCTCAGACCTGAATGAAGGAAATGAATTAATTGACATCC	413
Qy	241	AGATACCTGGCTAAATGCGGATTTTCGCAAAAGAGATGTGTCCCTGACCATAGAGATGTG	300
Db	414	AGATACCTGGCTAAATGCGGATTTTCGCAAAAGAGATGTGTCCCTGACCATAGAGATGTG	473
Qy	301	ACTCTAGCAGACAGTGGGATCTACTGTCTGCGCGGATCCAAATCCACAGCATTAATGAATAT	360
Db	474	ACTCTAGCAGACAGTGGGATCTACTGTCTGCGCGGATCCAAATCCACAGCATTAATGAATAT	533
Qy	361	GAAGAAATTTAACTCGAAGTGTGTCATCAAAACAGGCAAG	399
Db	534	GAAGAAATTTAACTCGAAGTGTGTCATCAAAACAGGCAAG	572

RESULT	LOCUS	DEFINITION
12 BG545730	BG545730	837 bp mRNA linear EST 04-APR-2001 605257003p1 NIH_MGC_77 Homo sapiens CDNA clone IMAGE:4701288 5', mRNA sequence.

ACCESSION	BG545730
VERSION	BG545730.1
KEYWORDS	GI:13544395 EST.

SOURCE ORGANISM	Homo sapiens (human) Homo sapiens
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

**REFERENCE**  
1 (bases 1 to 837)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC).  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
plate: L1CM1536 row: e column: 01  
High quality sequence stop: 722.  
Location/Qualifiers  
1..837

```

/organism="Homo sapiens"
/mol type="mRNA"
-db_xref="taxon:9606"
/clone="IMAGE:4701288"
/lab host="DH10B (T1 phage-resistant)"
/clone_1ib="NH MGC 77"
/ncore="Organ: lung; Vector: pMDR-LIB (Clontech); Site_1:
SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence
5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech laboratories (Palo Alto,

```

CA) . Note: this is a NIH MGC Library."

BASE COUNT	205 a	188 c	216 g	228 t
ORIGIN				

Query Match	43.9%	Score 396.6	DB 10	Length 837
Best Local Similarity	97.8%	Pred. No. 1.2e-108		
Matches 402; Conservative	0	Mismatches 5	Indels 0	Gaps 0

QY 1 TNGTTTCACACCTTCCCTTTGACGTCGTCCCGTCGCTGCTGCTGCTACTACTTACAAG 60

Db 111 ATGTTTTACACCTTCCCTTTGACTGTGTCCTGCTGCTGCTGCTGCTGCTACTACTTACAAG 170

QY 61 TCCTCAGAGTGAATACAGAGCGGAGTCCGTTCAGAACTGCTATCTGCTGCTTCTTAC 120

Db 171 TCCTCAGAGTGAATACAGAGCGGAGTCCGTTCAGAACTGCTATCTGCTGCTTCTTAC 230

QY 121 ACCCAGCCGCCCCAGGGAACCTCGTGCCCGTCTGCTGGGGCAAGAGACCTGTCTGTG 180

Db 211 ACCCAGCCGCCCCAGGGAACCTCGTGCCCGTCTGCTGGGGCAAGAGACCTGTCTGTG 290

QY 181 TTGGAATGTGGCAACGTGTGCTCAGAGCTGATGAAAGGATGTGAAATTTAGGACATCC 240

Db 281 TTGGAATGTGGCAACGTGTGCTCAGAGCTGATGAAAGGATGTGAAATTTAGGACATCC 350

QY 241 AGATACCTGGCTAAATATGGGGATTTCCGCAAGAGAGATGTGTCCCTGACCAATAGGAATGTG 300

Db 351 AGATACCTGGCTAAATATGGGGATTTCCGCAAGAGAGATGTGTCCCTGACCAATAGGAATGTG 410

QY 301 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGGATCCAAATCCCAAGCATTAATGAATGAT 360

Db 411 ACTCTAGCAGACAGTGGGATCTACTGCTGCCGGATCCAAATCCCAAGGCATTAATGAATGAT 470

QY 361 GAAAAATTTAACTCTGAGTGGTATCTCAAAACAGCGCAAGGTCAACCCCTTCA 411

Db 471 GAAAAATTTAACTCTGAGTGGTATCTCAAAACAGGTGAATGACATTTTCA 521

RESULT 13  
BI834671

LOCUS	519 bp	mrna	linear	EST 04-OCT-2001
DEFINITION	B1834671	603090448F1 NIH_MGC_120	Homo sapiens	cdna clone IMAGE:5229539 5',
	mrna sequence.			

ACCESSION	BI0834671
VERSION	BI0834671.1
KEYWORDS	GI:15946208
SOURCE	EST.
	Homo sapiens (human)

ORGANISM	REFERENCE
Homo sapiens	1 (pages 1 to 519)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC).  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LLM11577 row: 0 column: 12  
High quality sequence stop: 496.

```

FEATURES
source
location/Qualifiers
1..519
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5229539"
/lab_name="DH10B"
/clone_1fb="NH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector:

```

pcmv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 117 a 127 c 142 g 133 t

ORIGIN

Query Match 38.7%; Score 349.8; DB 12; Length 519;  
Best Local Similarity 99.4%; Pred. No. 1.4e-94;  
Matches 351; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTTTTCACATCTTCCCTTTGACTGTGCTCTGCTGCTGCTACTTACAAAG 60  
DB 145 ATGTTTTCACATCTTCCCTTTGACTGTGCTCTGCTGCTGCTACTTACAAAG 204  
QY 61 TCCTCAGAAATGGAATTCAGAGCGGAGTGGTTCAGAAATGCTATCTGCTCTTCTAC 120  
DB 205 TCCTCAGAAATGGAATTCAGAGCGGAGTGGTTCAGAAATGCTATCTGCTCTTCTAC 264  
QY 121 ACCCGAGCGCGCCGAGGGAACCTCGTCCGCTGCTGCGGCAAGAGCGCTGCTGTG 180  
DB 265 ACCCGAGCGCGCCGAGGGAACCTCGTCCGCTGCTGCGGCAAGAGCGCTGCTGTG 324  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTTGAATCC 240  
DB 325 TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTTGAATCC 384  
QY 241 AGATACGTGCTAATGCGGATTTCCGCAAGAGATGTGCTGACCATAGAAATG 300  
DB 385 AGATACGTGCTAATGCGGATTTCCGCAAGAGATGTGCTGACCATAGAAATG 444  
QY 301 ACTCTAGCAGACAGTGGATCTACTGCTGCGGATCCAAATCCAGGCAATAT 353  
DB 445 ACTCTAGCAGACAGTGGATCTACTGCTGCGGATCCAAATCCAGGCAATAT 497

RESULT 14  
LOCUS B0278485 1370 bp mRNA linear EST 07-MAY-2002  
DEFINITION AGENCOURT\_6773324.NIH\_MGC\_128 Homo sapiens cDNA clone IMAGE:5811526  
5', mRNA sequence.  
ACCESSION B0278485  
VERSION B0278485.1 GI:20488693  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1370)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: NCI  
CDNA Library Preparation: Michael Brownstein Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LICM2060 row: h column: 23  
High quality sequence stop: 183.  
Location/Qualifiers

FEATURES  
source  
1..1370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5811526"

/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phase-resistant)"  
/clone\_lib="NIH\_MGC\_128"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggc); Site 2: SfiI (ggccgctccggc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 2.2%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
5'-ATGAGTGGTATCAACGAGAGTGGGCTTATGAGCGCGG-3' and  
5'-ATCTTAGAGCGCGAGCGCGCCACATG-df(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH\_MGC 126 and NIH\_MGC 127). Library created in the laboratory of T. Ueda, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC Library."

BASE COUNT 277 a 486 c 299 g 308 t

ORIGIN

Query Match 38.2%; Score 345; DB 13; Length 1370;  
Best Local Similarity 95.4%; Pred. No. 7.3e-93;  
Matches 377; Conservative 0; Mismatches 15; Indels 3; Gaps 2;

QY 1 ATGTTTTCACATCTTCCCTTTGACTGTGCTCTGCTGCTGCTACTTACAAAG 60  
DB 113 ATGTTTTCACATCTTCCCTTTGACTGTGCTCTGCTGCTGCTACTTACAAAG 172  
QY 61 TCCTCAGAAATGGAATTCAGAGCGGAGTGGTTCAGAAATGCTATCTGCTCTTCTAC 120  
DB 173 TCCTCAGAAATGGAATTCAGAGCGGAGTGGTTCAGAAATGCTATCTGCTCTTCTAC 232  
QY 121 ACCCGAGCGCGCCGAGGGAACCTCGTCCGCTGCTGCGGCAAGAGCGCTGCTGTG 180  
DB 233 ACCCGAGCGCGCCGAGGGAACCTCGTCCGCTGCTGCGGCAAGAGCGCTGCTGTG 292  
QY 181 TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTTGAATCC 240  
DB 293 TTTGAATGTGGCAACGTGTGCTCAGACTGATGAAAGGATGTGAATTTTGAATCC 352  
QY 241 AGATACGTGCTAATGCGGATTTCCGCAAGAGATGTGCTGACCATAGAAATG 299  
DB 353 AGATACGTGCTAATGCGGATTTCCGCAAGAGATGTGCTGACCATAGAAATG 412  
QY 300 GACTCTAGCAGACAGTGGATCTACTGCTGCGGATCCAAATCCAGGCAATATGAATGA 359  
DB 413 GACTCTAGCAGACAGTGGATCTACTGCTGCGGATCCAAATCCAGGCAATATGAATGA 472  
QY 360 TGAATAATTAA--CTTGAAGTTGCTCATCAAC 392  
DB 473 TGAATAATTAACTTCCCGCAAGTGGTCAATCAAAAC 507

RESULT 15  
LOCUS BY741594 633 bp mRNA linear EST 17-DEC-2002  
DEFINITION BY741594 RIKEN full-length enriched, adult pancreas islet cells Mus musculus cDNA clone C820007812 5', mRNA sequence.  
ACCESSION BY741594  
VERSION BY741594.1 GI:27166585  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 633)  
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, T., Bono, H., Kondo, S., Nishida, I., Ose, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,



Mon Nov 24 13:30:28 2003

Job time : 1981.2 secs

us-10-004-633-37.rst

Page 12

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:56:41 ; Search time 10120.7 Seconds  
(without alignments)  
10954.308 Million cell updates/sec

Title: US-10-004-633-8  
Perfect score: 2710  
Sequence: 1 ngcgcaccacgcgcgcga.....aaaaaaaaagggcgccgc 2710

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb Da: \*  
2: gb Hcg: \*  
3: gb In: \*  
4: gb Om: \*  
5: gb Ov: \*  
6: gb Pat: \*  
7: gb Ph: \*  
8: gb Pl: \*  
9: gb Pr: \*  
10: gb Ro: \*  
11: gb Str: \*  
12: gb Sy: \*  
13: gb Un: \*  
14: gb Vi: \*  
15: em Ba: \*  
16: em Fun: \*  
17: em Hum: \*  
18: em In: \*  
19: em Mu: \*  
20: em Om: \*  
21: em Or: \*  
22: em Ov: \*  
23: em Pat: \*  
24: em Ph: \*  
25: em Pl: \*  
26: em Ro: \*  
27: em Str: \*  
28: em Un: \*  
29: em Vi: \*  
30: em Hcg\_hum: \*  
31: em Hcg\_inv: \*  
32: em Hcg\_other: \*  
33: em Hcg\_mus: \*  
34: em Hcg\_pln: \*  
35: em Hcg\_rtd: \*  
36: em Hcg\_mam: \*  
37: em Hcg\_vrt: \*  
38: em Sy: \*  
39: em Htgo\_hum: \*  
40: em Htgo\_mus: \*  
41: em Htgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	6	AR143561 Sequence
2	2706.6	99.9	2710	6	AR168930 Sequence
3	2706.6	99.9	2710	6	AR232678 Sequence
4	2706.6	99.9	2710	6	189415 Sequence 8
5	2624.6	96.8	2725	10	AF450241 Mus muscu
6	1886.6	69.6	223030	10	AL669948 Mouse DNA
7	1144.8	42.2	177087	2	AC119505
8	1144.8	42.2	233370	2	AC098539
9	844.4	31.2	862	10	AF39831 Mus muscu
10	843	31.1	843	6	AR143567 Sequence
11	843	31.1	843	6	AR168936 Sequence
12	843	31.1	843	6	AR232684 Sequence
13	499.8	18.4	1772	6	BD157433 Primer fo
14	499.8	18.4	1772	9	AK027334 Homo sapi
15	499.8	18.4	2236	6	AR143568 Sequence
16	499.8	18.4	2236	6	AR168937 Sequence
17	499.8	18.4	2236	6	AR232685 Sequence
18	498.2	18.4	2320	9	AF251707 Homo sapi
19	485.8	17.9	1116	9	AF450242
20	485.8	17.9	1116	9	AF450243 Homo sapi
21	456	16.8	903	6	AR143581 Sequence
22	456	16.8	903	6	AR168950 Sequence
23	456	16.8	903	6	AR232698 Sequence
24	456	16.8	906	6	AY069944 Homo sapi
25	336.4	12.4	769	6	BD146047
26	263.4	9.7	529	6	BD060055 Secretd
27	255.6	9.4	177087	2	AC119505 Rattus no
28	255.6	9.4	233370	2	AC098539 Rattus no
29	231.6	8.5	1012	9	BC020843 Homo sapi
30	219.8	8.1	438	6	BD076086 5' EST of
31	208.8	7.7	223030	10	AL669948 Mouse DNA
32	196.2	7.2	228827	2	AC135694 Rattus no
33	189.4	7.0	54375	2	AC011377 Homo sapi
34	189.4	7.0	54375	2	AC108587 Rattus no
35	104.2	3.8	213613	2	AC110244 Mus muscu
36	103.6	3.8	269920	2	AC094262 Rattus no
37	102.6	3.8	175542	2	AC112760 Rattus no
38	102.6	3.8	211694	2	AC141141 Rattus no
39	102	3.8	216060	2	AC097235 Rattus no
40	100	3.7	190756	2	AC107854 Mus muscu
41	96.8	3.6	257333	2	AC132722 Rattus no
42	96.4	3.6	236117	2	AC106381 Rattus no
43	95.2	3.5	203206	2	AC109866 Rattus no
44	94.4	3.5	188016	2	AC121256 Mus muscu
45	92.4	3.4	76318	2	AC100517 Mus muscu

#### ALIGNMENTS

RESULT 1  
AR143561  
LOCUS  
DEFINITION Sequence 8 from patent US 6204371.  
ACCESSION AR143561  
VERSION AR143561.1 GI:15104847  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 2710)  
AUTHORS Lewinson,D.Adam.  
TITLES Compositions and methods for the treatment and diagnosis of immune disorders  
JOURNAL Patent: US 6204371-A 8 20-MAR-2001;



FEATURES Location/Qualifiers  
source 1..2710  
/organism="unknown"  
BASE COUNT 789 a 592 c 644 g 678 t 7 others  
ORIGIN

Query Match 99.9%; Score 2706.6; DB 6; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGACCCACAGCGTCCGGAATTCCTCCCAAGTACTCATGTTTCAGGCTTACCTCA 61  
DB 2 GTGACCCACAGCGTCCGGAATTCCTCCCAAGTACTCATGTTTCAGGCTTACCTCA 61  
QY 62 ACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121  
DB 62 ACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121  
QY 122 TTGAGGTTGGTAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181  
DB 122 TTGAGGTTGGTAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181  
QY 182 TTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
DB 182 TTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
QY 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGAGATACCAAGTAAAGGCG 301  
DB 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGAGATACCAAGTAAAGGCG 301  
QY 302 ATCTCAACAAAGAGATGTGTCTCTGATCATATAAGATGATGATGATGATGATGATGAT 361  
DB 302 ATCTCAACAAAGAGATGTGTCTCTGATCATATAAGATGATGATGATGATGATGATGAT 361  
QY 362 CCTACTGCTGACAGATACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
DB 362 CCTACTGCTGACAGATACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
QY 422 TTGACATCAAGAGCCCAAGGTCATCTCAGCTCAGACTGCGCCATGCGGACTTACTACG 481  
DB 422 TTGACATCAAGAGCCCAAGGTCATCTCAGCTCAGACTGCGCCATGCGGACTTACTACG 481  
QY 482 CTCTCCCAAGAACCTTCAACCAAGAGAGAAATGCTTCAGAGACACAGACCTGCTGACCC 541  
DB 482 CTCTCCCAAGAACCTTCAACCAAGAGAGAAATGCTTCAGAGACACAGACCTGCTGACCC 541  
QY 542 TCCATATATACAAATGGAACAAAATTTCCATGAGGCTGATGAAATTTAAGACTCTGAG 601  
DB 542 TCCATATATACAAATGGAACAAAATTTCCATGAGGCTGATGAAATTTAAGACTCTGAG 601  
QY 602 AAAAGATCAGAACTGCTATCCATGAGTGGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 661  
DB 602 AAAAGATCAGAACTGCTATCCATGAGTGGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 661  
QY 662 TTATCATGCTGCTTCAATCTTAAATGATCTCTGTAAGAAAAGAAATTTACGAGTT 721  
DB 662 TTATCATGCTGCTTCAATCTTAAATGATCTCTGTAAGAAAAGAAATTTACGAGTT 721  
QY 722 TGAGCCTTATTAACACTGCGCAACTTGTCTCAGAGAGGTTGGCAAAATGACAGAGCTCA 781  
DB 722 TGAGCCTTATTAACACTGCGCAACTTGTCTCAGAGAGGTTGGCAAAATGACAGAGCTCA 781  
QY 782 GGATTTGGCTTGAAGAAATATATCTACACATCGAGAGAAAGTATATGAAATGAGAGATT 841  
DB 782 GGATTTGGCTTGAAGAAATATATCTACACATCGAGAGAAAGTATATGAAATGAGAGATT 841  
QY 842 CAATATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901  
DB 842 CAATATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901  
QY 902 ACTTTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961  
DB 902 ACTTTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961

QY 962 ATATGTCACTTGGCAACTCATTTGAGAGGTTCTGACCAAGCCCATGAGAAAAGATTCCA 1021  
DB 962 ATATGTCACTTGGCAACTCATTTGAGAGGTTCTGACCAAGCCCATGAGAAAAGATTCCA 1021  
QY 1022 GTTTTCTGGGGATTAATTAATCAACAAGGGGATTCAGTGTACTCATGCTTACATTTGAAT 1081  
DB 1022 GTTTTCTGGGGATTAATTAATCAACAAGGGGATTCAGTGTACTCATGCTTACATTTGAAT 1081  
QY 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTGGATCTGCCATCTCCAGAGACTTCAATCATG 1141  
DB 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTGGATCTGCCATCTCCAGAGACTTCAATCATG 1141  
QY 1142 CGTGTGAAGCTCACTGCTGCTTATCATATTAAGAAATGTTATGTTATGTTGTTGAGA 1201  
DB 1142 CGTGTGAAGCTCACTGCTGCTTATCATATTAAGAAATGTTATGTTATGTTGTTGAGA 1201  
QY 1202 CATAGAGTTTGTGGATATCCGCAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGGCTTA 1261  
DB 1202 CATAGAGTTTGTGGATATCCGCAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGGCTTA 1261  
QY 1262 GATAGGAAGTCCGGCTTTTGTGATGTTGAAAAATCTTAAAGAAATTTGATGCTTTCT 1321  
DB 1262 GATAGGAAGTCCGGCTTTTGTGATGTTGAAAAATCTTAAAGAAATTTGATGCTTTCT 1321  
QY 1322 AGAGATTTTGTGACTTTGAAAGATTAAGAAAAGCCAGGGTGGATATGCTTAAACAGATAT 1381  
DB 1322 AGAGATTTTGTGACTTTGAAAGATTAAGAAAAGCCAGGGTGGATATGCTTAAACAGATAT 1381  
QY 1382 AACTTGGGAACCTTAGGCGAGAGGGTATTAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441  
DB 1382 AACTTGGGAACCTTAGGCGAGAGGGTATTAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441  
QY 1442 AGACTGTCTCAGCATTCGAAAGCGAAATTAACATPAGACAGAGAGGCTGAGATGA 1501  
DB 1442 AGACTGTCTCAGCATTCGAAAGCGAAATTAACATPAGACAGAGAGGCTGAGATGA 1501  
QY 1502 GGCTGGAGAGTGAAGTGTATGTTGACAGACAGAGAAATTTATTTTGTGATCTGAGAC 1561  
DB 1502 GGCTGGAGAGTGAAGTGTATGTTGACAGACAGAGAAATTTATTTTGTGATCTGAGAC 1561  
QY 1562 CCACATGAAAAGCTTAGGCTGCTGAGAGATCTTGTAGACTCAAGAGATGAGAGATGA 1621  
DB 1562 CCACATGAAAAGCTTAGGCTGCTGAGAGATCTTGTAGACTCAAGAGATGAGAGATGA 1621  
QY 1622 AGGCAACAGATTCCTGGGGCTTGGCTGCACTGACCTTAAGCTTGAAGTGTGCTCA 1681  
DB 1622 AGGCAACAGATTCCTGGGGCTTGGCTGCACTGACCTTAAGCTTGAAGTGTGCTCA 1681  
QY 1682 GTCCCAAGAGTCCCTGCTCAGAGATGAATGAGTATCTGGCGGATGCTCATGAG 174  
DB 1682 GTCCCAAGAGTCCCTGCTCAGAGATGAATGAGTATCTGGCGGATGCTCATGAG 174  
QY 1742 GGTGTCTCTCTCTCTCAGAGAGATGATGAGTGGAGTGTGAGAGTGTGAGAGTGTGAG 1801  
DB 1742 GGTGTCTCTCTCTCTCAGAGAGATGATGAGTGGAGTGTGAGAGTGTGAGAGTGTGAG 1801  
QY 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGAT 1861  
DB 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGAT 1861  
QY 1862 TCTCTATACATGATATCTTACAGAGCTCTCTGCTGCTCTGTTAAGACATGAGTGGAG 1921  
DB 1862 TCTCTATACATGATATCTTACAGAGCTCTCTGCTGCTCTGTTAAGACATGAGTGGAG 1921  
QY 1922 CATGGCAGAGAGTCCAGTAAATTTATCCAGACTCAGAGAGGCTGAGAGAGAGCTGGA 1981  
DB 1922 CATGGCAGAGAGTCCAGTAAATTTATCCAGACTCAGAGAGGCTGAGAGAGAGCTGGA 1981  
QY 1982 GAGTTGAGAGAGCTGTGCTCAACATGCTCAAGCTTTCTTACACAGAAAAGGTTACC 2041  
DB 1982 GAGTTGAGAGAGCTGTGCTCAACATGCTCAAGCTTTCTTACACAGAAAAGGTTACC 2041



QY 2042 CGCAGCAGCCTGCTGTCTGTAAAGAAACCTGCGAAGGCAAACTTTGACTGTGTG 2101  
 Db 2042 CGCAGCAGCCTGCTGTCTGTAAAGAAACCTGCGAAGGCAAACTTTGACTGTGTG 2101  
 QY 2102 TGTCTCAAGGGAACTGACTGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTC 2161  
 Db 2102 TGTCTCAAGGGAACTGACTGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTC 2161  
 QY 2162 TGACAGAAACAACCCGGTGACTGGGACATACGAAAGGAGAGCTTTGACAGCAATCTATA 2221  
 Db 2162 TGACAGAAACAACCCGGTGACTGGGACATACGAAAGGAGAGCTTTGACAGCAATCTATA 2221  
 QY 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTCGTACCAAAATGATTTTCAAGCCGGTGG 2281  
 Db 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTCGTACCAAAATGATTTTCAAGCCGGTGG 2281  
 QY 2282 ACCCTCAGTTTATCTGGCTTACAGCTGCTGCCAGTGTCTGATCTGTGCTGCTCC 2341  
 Db 2282 ACCCTCAGTTTATCTGGCTTACAGCTGCTGCCAGTGTCTGATCTGTGCTGCTCC 2341  
 QY 2342 ATCTATAACAGATCAATTAATTAAGACCCGAGTGAATAATTAAGTACAGCAAAAGT 2401  
 Db 2342 ATCTATAACAGATCAATTAATTAAGACCCGAGTGAATAATTAAGTACAGCAAAAGT 2401  
 QY 2402 AGCTTTGTTCAAGATTTTTTGTGCAATGGGAGCAACTGTGTACATCAGAGCACTGT 2461  
 Db 2402 AGCTTTGTTCAAGATTTTTTGTGCAATGGGAGCAACTGTGTACATCAGAGCACTGT 2461  
 QY 2462 TAGTGAGGACACCAAAACCTGTGTACCGTTTTCATGTATGAATTTTGTGTTAGT 2521  
 Db 2462 TAGTGAGGACACCAAAACCTGTGTACCGTTTTCATGTATGAATTTTGTGTTAGT 2521  
 QY 2522 TGCCTCTAGAGCTGCTGAGAGTCTGCTGCTTTCTTAGGTGGTATGAAAGGAGCAATC 2581  
 Db 2522 TGCCTCTAGAGCTGCTGAGAGTCTGCTGCTTTCTTAGGTGGTATGAAAGGAGCAATC 2581  
 QY 2582 TAACAAAATCCATTAGAGATTAACAGCTCTCATGCGAAGGAGAAACTATCTCAATGT 2641  
 Db 2582 TAACAAAATCCATTAGAGATTAACAGCTCTCATGCGAAGGAGAAACTATCTCAATGT 2641  
 QY 2642 TTAAAGTATTAATACTGACTGGCAAGTCTTTGAGCATTAATAAAAAAAAAAAAAAG 2701  
 Db 2642 TTAAAGTATTAATACTGACTGGCAAGTCTTTGAGCATTAATAAAAAAAAAAAAAAG 2701  
 QY 2702 GCGCGCGCG 2710  
 Db 2702 GCGCGCGCG 2710

RESULT 2  
 ARI68930  
 LOCUS ARI68930 2710 bp DNA linear PAT 17-DEC-2001  
 DEFINITION Sequence 8 from patent US 6288218.  
 ACCESSION ARI68930  
 VERSION ARI68930.1 GI:17905107  
 KEYWORDS  
 SOURCE unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 2710)  
 AUTHORS Levynson,D.Adam.  
 TITLE Compositions and methods for the treatment and diagnosis of immune disorders  
 JOURNAL Patent: US 6288218-A 8 11-SEP-2001;  
 FEATURES  
 source 1. 2710  
 BASE COUNT 789 a 592 c 644 g 678 t 7 others  
 ORIGIN

Query Match 99.9%; Score 2706.6; DB 6; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGACCAAGCGCTCCGGAATTTCCCTCCCAAGTACTCATGTTTCAAGTCTTACCTCA 61  
 Db 2 GTGACCAAGCGCTCCGGAATTTCCCTCCCAAGTACTCATGTTTCAAGTCTTACCTCA 61  
 QY 62 ACTGTCTCTGCTGCTGCACTACTTGTGCAAGGTCAATGGAAGTGTATTAAG 121  
 Db 62 ACTGTCTCTGCTGCTGCACTACTTGTGCAAGGTCAATGGAAGTGTATTAAG 121  
 QY 122 TTGAGTTGGTAAAGGCTATCTGCTCCGACATTAACCTTACCTTACCTGGAAC 181  
 Db 122 TTGAGTTGGTAAAGGCTATCTGCTCCGACATTAACCTTACCTTACCTGGAAC 181  
 QY 182 TTGTCCTATGCTGCTGCTGCAAGGATTTCTGTCTTGTGCACTGATCAATGATG 241  
 Db 182 TTGTCCTATGCTGCTGCTGCAAGGATTTCTGTCTTGTGCACTGATCAATGATG 241  
 QY 242 TCAGAACTGATGAAGAATGTGACATATCAAGAAATCCAGAGATACAGCTTAAAGG 301  
 Db 242 TCAGAACTGATGAAGAATGTGACATATCAAGAAATCCAGAGATACAGCTTAAAGG 301  
 QY 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTACTTGTGATCAATGGA 361  
 Db 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTACTTGTGATCAATGGA 361  
 QY 362 CCTACTGCTGAGGATACAGTTCCCTGCTTATGAATGATTAATAATTAAGCAAT 421  
 Db 362 CCTACTGCTGAGGATACAGTTCCCTGCTTATGAATGATTAATAATTAAGCAAT 421  
 QY 422 TAGACATCAAGCAGCAGGCTCACTCCAGCTCAGACTCCCATGAGGACTCTACTAG 481  
 Db 422 TAGACATCAAGCAGCAGGCTCACTCCAGCTCAGACTCCCATGAGGACTCTACTAG 481  
 QY 482 CTCTCTCAAGAACCTTAAACCAAGGAGAAATGTGTTAGAGACACACACTGTGTACC 541  
 Db 482 CTCTCTCAAGAACCTTAAACCAAGGAGAAATGTGTTAGAGACACACACTGTGTACC 541  
 QY 542 TCCATTAATTAACATGGAACAAATAATTTCCATGAGGCTGATGAATTAAGACTGAG 601  
 Db 542 TCCATTAATTAACATGGAACAAATAATTTCCATGAGGCTGATGAATTAAGACTGAG 601  
 QY 602 AAACGATCAGAACTGCTATCCATATGAGAGTCTGCTGAGTGAACCTGAGAC 661  
 Db 602 AAACGATCAGAACTGCTATCCATATGAGAGTCTGCTGAGTGAACCTGAGAC 661  
 QY 662 TTATCATTTGCTCTTAATCTTAAATGTATTTCTGTAAGAAAGAAAGTTATCGAGT 721  
 Db 662 TTATCATTTGCTCTTAATCTTAAATGTATTTCTGTAAGAAAGAAAGTTATCGAGT 721  
 QY 722 TGAGCTTATTAACCTGCGCAACTGCTCCAGAGAGGTTGGCAATGACAGAGACATCA 781  
 Db 722 TGAGCTTATTAACCTGCGCAACTGCTCCAGAGAGGTTGGCAATGACAGAGACATCA 781  
 QY 782 GGATTCGCTGAGAGAAATATCTTACACATCGAGAGAACTATATGAGTGAATTT 841  
 Db 782 GGATTCGCTGAGAGAAATATCTTACACATCGAGAGAACTATATGAGTGAATTT 841  
 QY 842 CAAATGATTAATCTGCTAGCTCAACAGCAGCAGCACTCTGACCTGACCTGCC 901  
 Db 842 CAAATGATTAATCTGCTAGCTCAACAGCAGCAGCACTCTGACCTGACCTGCC 901  
 QY 902 ACTTTTAAAGGCTGCTGCTTATCTGACTTGTGATTTTCCCTTTTGGAAAACTATG 961  
 Db 902 ACTTTTAAAGGCTGCTGCTTATCTGACTTGTGATTTTCCCTTTTGGAAAACTATG 961  
 QY 962 ATATGTACTTGGCACTCATTTGAGAGTTTGAACACAGCACTGAGAAAAAGTTCCA 1021  
 Db 962 ATATGTACTTGGCACTCATTTGAGAGTTTGAACACAGCACTGAGAAAAAGTTCCA 1021  
 QY 1022 GTTTTCTGGGGATTAATTAATCAAGAGGATTTGACGTGATCTGCTCATTTGAAT 1081  
 Db 1022 GTTTTCTGGGGATTAATTAATCAAGAGGATTTGACGTGATCTGCTCATTTGAAT 1081

QY	1082	GCTCATTATTAACCCCTGAGTTTCAGGAGTCGGATCTCCACCTCCAGACCTTCAATCAG	1141
Db	1082	GCTCATTATTAACCCCTGAGTTTCAGGAGTCGGATCTCCACCTCCAGACCTTCAATCAG	1141
QY	1142	CGTGTGAAGCTCACTCGTCTTTTCATATTAGGAATGTAGTGTATGTCCTTTGAGA	1201
Db	1142	CGTGTGAAGCTCACTCGTCTTTTCATATTAGGAATGTAGTGTATGTCCTTTGAGA	1201
QY	1202	CATAGAGTTTGTGTATATCCGCAAGCTCTGAAACAGTATAGGGGAAATTAAGGGCTTA	1261
Db	1202	CATAGAGTTTGTGTATATCCGCAAGCTCTGAAACAGTATAGGGGAAATTAAGGGCTTA	1261
QY	1262	GATTAGGAAGTCGCGTCTTTGTGATGTGGAAAATCTTAAAGAAAGTTGGAGCTTTCT	1321
Db	1262	GATTAGGAAGTCGCGTCTTTGTGATGTGGAAAATCTTAAAGAAAGTTGGAGCTTTCT	1321
QY	1322	AGAGATTTCTGACTTTGAAAGATTAAAGAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
Db	1322	AGAGATTTCTGACTTTGAAAGATTAAAGAAAAGCCAGGTGGCATATGCTTAAACGATAT	1381
QY	1382	AACTTGGGAACCTTAGGCGAGGGGTGATAGTTCAAGTCAAGCCAGGGCTATGCTGTA	1441
Db	1382	AACTTGGGAACCTTAGGCGAGGGGTGATAGTTCAAGTCAAGCCAGGGCTATGCTGTA	1441
QY	1442	AGACTGTCTCMACATCCAAAGCGAAATTAACATAGAACACGACGAGGCTGGAGATGA	1501
Db	1442	AGACTGTCTCMACATCCAAAGCGAAATTAACATAGAACACGACGAGGCTGGAGATGA	1501
QY	1502	GGCTTCGACAGTGAAGGTGATGTGTACACGACGAGGAATCTAATTATTGATCGTAGAC	1561
Db	1502	GGCTTCGACAGTGAAGGTGATGTGTACACGACGAGGAATCTAATTATTGATCGTAGAC	1561
QY	1562	CCACATGAAAAAGCTAGGCTGTGTAGAGCATCTTGTAGACTCAAGAGATGAGAGGTAA	1621
Db	1562	CCACATGAAAAAGCTAGGCTGTGTAGAGCATCTTGTAGACTCAAGAGATGAGAGGTAA	1621
QY	1622	AGGCACAACAGTATCCCGGGGCTTCGTCGATCAGCTTAAGCTAGTGGCTGAGTTCCAA	1681
Db	1622	AGGCACAACAGTATCCCGGGGCTTCGTCGATCAGCTTAAGCTAGTGGCTGAGTTCCAA	1681
QY	1682	GTCACAAGAGTCCCTGTCTCMAGTAAAGATGATCTAGATATCTGGCCATGTCATGAG	1741
Db	1682	GTCACAAGAGTCCCTGTCTCMAGTAAAGATGATCTAGATATCTGGCCATGTCATGAG	1741
QY	1742	GGTTGTCTCTCTCTCAGAAAGACATGCAATGWCCTTGACAACACACACACACAC	1801
Db	1742	GGTTGTCTCTCTCTCAGAAAGACATGCAATGWCCTTGACAACACACACACACAC	1801
QY	1802	ACACACACACACACACACACACACACATGMAATGAAGTCTCTGTGTCTCTAC	1861
Db	1802	ACACACACACACACACACACACACACATGMAATGAAGTCTCTGTGTCTCTAC	1861
QY	1862	TCTCTAATAATGATATCTCTAAGAGCTCTCTGCTCTGTGTAAGACATGAGTGGAG	1921
Db	1862	TCTCTAATAATGATATCTCTAAGAGCTCTCTGCTCTGTGTAAGACATGAGTGGAG	1921
QY	1922	CATGCGAAGGCAAGTCAAGTAAATTTATCCAGCACTCAGAAAGCTGGACAGAGAGTGA	1981
Db	1922	CATGCGAAGGCAAGTCAAGTAAATTTATCCAGCACTCAGAAAGCTGGACAGAGAGTGA	1981
QY	1982	GAGTTTCAGAGACATGTGCTCCAAACCTGCACAACTCTTCTTAACAAGAAAAGGTTAAC	2041
Db	1982	GAGTTTCAGAGACATGTGCTCCAAACCTGCACAACTCTTCTTACAAGAAAAGGTTAAC	2041
QY	2042	CGCAAGCAGCCGTCGTCTGTAAAGGAAACCTCGGAAAGGCAAACTTTGACTGTGTG	2101
Db	2042	CGCAAGCAGCCGTCGTCTGTAAAGGAAACCTCGGAAAGGCAAACTTTGACTGTGTG	2101
QY	2102	TGCTCAAGGGGAACTGACTCAACAACTTCTCAATTCCTGAGAGAAACTGAGAGCTTTTC	2161
Db	2102	TGCTCAAGGGGAACTGACTCAACAACTTCTCAATTCCTGAGAGAAACTGAGAGCTTTTC	2161
QY	2162	TGACAGAAAGCAACCGGTGACTGGGACATACAAAGGCAAGACTCTTGACGACATCTATA	2221

Db	2162	TGACAGAAAGAACCAACCGGTGACTGGGACATACGAAGGCAGAGCTCTGGACGAATCTATA	2221
OY	2222	TAGTCAGCAAAATATTTCTTTGGGAGGACAGTCGTCAACCAATTGATTTCCAGCCGGTGG	2281
Db	2222	TAGTCACCAAAATATTTCTTTGGGAGGACAGTCGTCAACCAATTGATTTCCAGCCGGTGG	2281
OY	2282	ACCTCAGTTTCATCTGGCTTACAGACTGCTGCCAGTGCCCTTGATCTGTGTGGCTCCC	2341
Db	2282	ACCTCAGTTTCATCTGGCTTACAGACTGCTGCCAGTGCCCTTGATCTGTGTGGCTCCC	2341
OY	2342	ATCTATAACGAATCAAAATTAATTAAGNACCCCGAGTGAATATTTAAGTGAAGCAGAAAGT	2401
Db	2342	ATCTATAACGAATCAAAATTAATTAAGNACCCCGAGTGAATATTTAAGTGAAGCAGAAAGT	2401
OY	2402	AGCTTTGTCAAAAGATTTTTTGATTTGGGAGGACATGTGTACATAGAGGACATCTGT	2461
Db	2402	AGCTTTGTCAAAAGATTTTTTGATTTGGGAGGACATGTGTACATAGAGGACATCTGT	2461
OY	2462	TAGTGAAGACACCAAAACCTGTGGTACCGTTTTTTTCATGTATGAATTTGTGTTTAGT	2521
Db	2462	TAGTGAAGACACCAAAACCTGTGGTACCGTTTTTTTCATGTATGAATTTGTGTTTAGT	2521
OY	2522	TGCTTTAGCTAGCTGTGGAGGTCCTGGCTTTTCTTAGTGGGTATGGAAGGAGACATC	2581
Db	2522	TGCTTTAGCTAGCTGTGGAGGTCCTGGCTTTTCTTAGTGGGTATGGAAGGAGACATC	2581
OY	2582	TAACAAAATCCATTAGAGATPAACAGCTCTCATGTCAGAAAGGGAAAACTAATCTCAAAATGTT	2641
Db	2582	TAACAAAATCCATTAGAGATPAACAGCTCTCATGTCAGAAAGGGAAAACTAATCTCAAAATGTT	2641
OY	2642	TTAAAGTAAATAACTGTACTGGCAAAAGTACTTTGACACATAAAAAATTTTTTTTTTAAAG	2701
Db	2642	TTAAAGTAAATAACTGTACTGGCAAAAGTACTTTGACACATAAAAAATTTTTTTTTTAAAG	2701
OY	2702	GGCGGCGCGC 2710	
Db	2702	GGCGGCGCGC 2710	

RESULT	3
LOCUS	AR232678
DEFINITION	Sequence 8 from patent US 6455685.
ACCESSION	AR232678
VERSION	AR232678.1 GI:27274955
KEYWORDS	.
SOURCE	. Unknown.
ORGANISM	. UNKNOWN.
REFERENCE	. Unclassified. 1 (bases 1 to 2710)
AUTHORS	Levinson,D.A.
TITLE	Compositions and methods for the treatment and diagnosis of immune disorders Patent: US 6455685-A 8-24-SEP-2002;
JOURNAL FEATURES	Location/Qualifiers . . 2710 /organism="unknown"
BASE COUNT	789 a . 592 c 644 g 678 t 7 others
ORIGIN	

Query Match	99.9%	Score 2706.6	DB 6	Length 2710
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2709	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	2	GTGCAGCCAGCGCGGATTTCCCTCCCGCAGACTCATGTTTACAGTCTTACCTCA	61	
Db	2	GTGCAGCCAGCGCGGATTTCCCTCCCGCAGACTCATGTTTACAGTCTTACCTCA	61	
QY	62	ACTGTGTCCTGCTCTGCTGCACCTACTACTTGCAGAGTCAATTGGAAGATGTTATAAG	121	
Db	62	ACTGTGTCCTGCTCTGCTGCACCTACTACTTGCAGAGTCAATTGGAAGATGTTATAAG	121	

QY 122 TTGAGTTGGTAAAAATGCTATCTGCCCCGAGTTACACTCTACTCATCTGGGACAC 181  
Db 122 TTGAGTTGGTAAAAATGCTATCTGCCCCGAGTTACACTCTACTCATCTGGGACAC 181  
QY 182 TTGTCCTATGTCCTGGGGCAAGGATTTCTCTTGGTCAACAGTACCAATGAGTTGC 241  
Db 182 TTGTCCTATGTCCTGGGGCAAGGATTTCTCTTGGTCAACAGTACCAATGAGTTGC 241  
QY 242 TCAGAACTGATGAAGAAATGATGACATATCAAGAAATCCAGACAGATACCAAGCTAAAGGGCG 301  
Db 242 TCAGAACTGATGAAGAAATGATGACATATCAAGAAATCCAGACAGATACCAAGCTAAAGGGCG 301  
QY 302 ATCTCAACAAAGAGATGTGTCTCTGATCATTAAGAAATGATGATCTGATGACATGGGA 361  
Db 302 ATCTCAACAAAGAGATGTGTCTCTGATCATTAAGAAATGATGATCTGATGACATGGGA 361  
QY 362 CTCTACTGCTGAGATACAGTTCCCTGTCTTATGATTAATAAATTTAGAACTGAAT 421  
Db 362 CTCTACTGCTGAGATACAGTTCCCTGTCTTATGATTAATAAATTTAGAACTGAAT 421  
QY 422 TAGACATAAAGCGCCAGAGGTCACTCCAGCTCAGACCTGCCATGGGACCTTACTACAG 481  
Db 422 TAGACATAAAGCGCCAGAGGTCACTCCAGCTCAGACCTGCCATGGGACCTTACTACAG 481  
QY 482 CTCTCCAAAGAACCTTAACACAGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541  
Db 482 CTCTCCAAAGAACCTTAACACAGAGAGAAATGTTTCAAGACACAGACACTGTGACCC 541  
QY 542 TCATATATTAACATGAACAAATAATTTCCACATGGGCTGATGAATTAAGACCTTGAG 601  
Db 542 TCATATATTAACATGAACAAATAATTTCCACATGGGCTGATGAATTAAGACCTTGAG 601  
QY 602 AAAAGATAGAACTGCTATCCACATTTGAGAGTGTCTGCTGGGTTGACCTCGGAC 661  
Db 602 AAAAGATAGAACTGCTATCCACATTTGAGAGTGTCTGCTGGGTTGACCTCGGAC 661  
QY 662 TTATCATTTGTTCTTAATCCCTTAATGATGATTCCTGTAGAGAAAGAGTTATGAGTT 721  
Db 662 TTATCATTTGTTCTTAATCCCTTAATGATGATTCCTGTAGAGAAAGAGTTATGAGTT 721  
QY 722 TGAAGCTTATTAACATGSCCAACTTGTCTCCAGAGAGGTTGGCAATGACAGACGATCA 781  
Db 722 TGAAGCTTATTAACATGSCCAACTTGTCTCCAGAGAGGTTGGCAATGACAGACGATCA 781  
QY 782 GGAATTCGCTGAGAGAAATATCTACACATCGAGAGAAAGTATATGAGTGGAGATT 841  
Db 782 GGAATTCGCTGAGAGAAATATCTACACATCGAGAGAAAGTATATGAGTGGAGATT 841  
QY 842 CAATGAGTACTACTGCTACGTCACACAGCAGACAGCCATCCGCTCGAGCTGCC 901  
Db 842 CAATGAGTACTACTGCTACGTCACACAGCAGACAGCCATCCGCTCGAGCTGCC 901  
QY 902 ACTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTCCCTTTKTAGAAAATCTATGTG 961  
Db 902 ACTTTAAAGGCTGCTTCAATTTCTGAATTTGATTTCCCTTTKTAGAAAATCTATGTG 961  
QY 962 ATATGTCATCTTGGCAACCTCATGTGAGGTTCTGACACACAGCCATCGAAGAAAGATTTCA 1021  
Db 962 ATATGTCATCTTGGCAACCTCATGTGAGGTTCTGACACACAGCCATCGAAGAAAGATTTCA 1021  
QY 1022 GTTTTCTGGGATATTAATTAACACAAAGGAGATTGACATGTAATCTCATGTAAT 1081  
Db 1022 GTTTTCTGGGATATTAATTAACACAAAGGAGATTGACATGTAATCTCATGTAAT 1081  
QY 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTCGATCTCCACTCCAGAGACTTCAATCAG 1141  
Db 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTCGATCTCCACTCCAGAGACTTCAATCAG 1141  
QY 1142 GGTGTTGAAGCTCACTGCTGCTTTCATATTAAGAAATGTTAGTGTCTTTTGA 1201  
Db 1142 GGTGTTGAAGCTCACTGCTGCTTTCATATTAAGAAATGTTAGTGTCTTTTGA 1201  
QY 1202 CATAGAGTTTGTGTATATCCGCAAGCTCTGAAAGGTAGGGGAAATAAGGGCTAA 1261

Db 1202 CATAGAGTTTGTGTATATCCGCAAGCTCTGAAACAGTATAGGGGAAATAAGGGCTAA 1261  
QY 1262 GATAGAAAGTCCGCTCTTTGTTGATGTTGAAAAATCTTAAGAAATTTGATGCTTTTCT 1321  
Db 1262 GATAGAAAGTCCGCTCTTTGTTGATGTTGAAAAATCTTAAGAAATTTGATGCTTTTCT 1321  
QY 1322 AGAGATTTCTGACCTGTAAGAAATTAAGAAAGAGAGTGCATATGCTTAACAGATAT 1381  
Db 1322 AGAGATTTCTGACCTGTAAGAAATTAAGAAAGAGAGTGCATATGCTTAACAGATAT 1381  
QY 1382 AACTTGGAGACCTTATAGCAGAGGTTGATTAATTAAGTCAAGCCAGGCTATGCTGTA 1441  
Db 1382 AACTTGGAGACCTTATAGCAGAGGTTGATTAATTAAGTCAAGCCAGGCTATGCTGTA 1441  
QY 1442 AGACTGTCTCAGTCATCCAAAGACAAATAATTAACATAGACAGCAGAGGCTGAGATGA 1501  
Db 1442 AGACTGTCTCAGTCATCCAAAGACAAATAATTAACATAGACAGCAGAGGCTGAGATGA 1501  
QY 1502 GGTCTCGACAGTGAAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
Db 1502 GGTCTCGACAGTGAAGTGCATTTGTTGTAACAGCAGAGAACTATATTTGATCTGAGAC 1561  
QY 1562 CCACATGAAAAAGCTAGGCTGTGATAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
Db 1562 CCACATGAAAAAGCTAGGCTGTGATAGCATGCTTGTAGACTCAAGAGATGAGAGTAA 1621  
QY 1622 AGGCACAAACATATCCCGGGGCTTGTGCTGAGTCAAGTCAAGTCTGAGTTCGAA 1681  
Db 1622 AGGCACAAACATATCCCGGGGCTTGTGCTGAGTCAAGTCAAGTCTGAGTTCGAA 1681  
QY 1682 GTCCAAAGAGTCCCTGTCTCAGAAAGTAAAGTGTGATGATGCTGAGTCCATGAG 1741  
Db 1682 GTCCAAAGAGTCCCTGTCTCAGAAAGTAAAGTGTGATGATGCTGAGTCCATGAG 1741  
QY 1742 GGTGTCTCTCTCTCAGAAAGACATGACATGCTGCTGACACACACACACACAC 1801  
Db 1742 GGTGTCTCTCTCTCAGAAAGACATGACATGCTGCTGACACACACACACACACAC 1801  
QY 1802 ACAACACACACACACACACACACACACACATGATGAGTTCTCTGCTGCTCTAC 1861  
Db 1802 ACAACACACACACACACACACACACACACATGATGAGTTCTCTGCTGCTCTAC 1861  
QY 1862 TCTCTATTAACATGATATCTCTACAGACCTCTCTGCTCTGTTAAAGCATGAGTGGAG 1921  
Db 1862 TCTCTATTAACATGATATCTCTACAGACCTCTCTGCTCTGTTAAAGCATGAGTGGAG 1921  
QY 1922 CATGGCAGAGCAGTCCAGTATTTATTCACACATCAGAAAGCTGGACAGAGGCTGGA 1981  
Db 1922 CATGGCAGAGCAGTCCAGTATTTATTCACACATCAGAAAGCTGGACAGAGGCTGGA 1981  
QY 1982 GAGTTTCAGAGACTGTGCCCAACACTGCTCTTCTTAACAAAGAAAGGTTAC 2041  
Db 1982 GAGTTTCAGAGACTGTGCCCAACACTGCTCTTCTTAACAAAGAAAGGTTAC 2041  
QY 2042 CGCAAGCAGCTGCTGTCTGTTAAAGAAACCTGCGCAAGAAAGCAATTTGATGTTG 2101  
Db 2042 CGCAAGCAGCTGCTGTCTGTTAAAGAAACCTGCGCAAGAAAGCAATTTGATGTTG 2101  
QY 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCCATCTCTGAGGAACTGAGCTGTTTC 2161  
Db 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCCATCTCTGAGGAACTGAGCTGTTTC 2161  
QY 2162 TGAAGAAAGCAACCGGTGACTGAGCATTAAGAGGAGAGCTTTGACAGCAATCTATA 2221  
Db 2162 TGAAGAAAGCAACCGGTGACTGAGCATTAAGAGGAGAGCTTTGACAGCAATCTATA 2221  
QY 2222 TAGTCAGAAATATTTCTTTGGAGAGCAGTGTCAACAAATTTGATTTTCAAGCCGTTG 2281  
Db 2222 TAGTCAGAAATATTTCTTTGGAGAGCAGTGTCAACAAATTTGATTTTCAAGCCGTTG 2281  
QY 2282 AACTCAGTTTATGCTGCTTACAGCTGCTGCCAGAGTCCCTGATCTGTGCTGCCCTCC 2341

Db 2282 ACCTAGTTTCATCTGCTTACAGCTGCTGCTGCCAGTCCCTTGATCTGTGCTGCTCC 2341

Qy 2342 ATCTATACAGAAATCAATTAATAGACCCCGAGTGAATAATTAAGTACAGAAAGT 2401

Db 2342 ATCTATACAGAAATCAATTAATAGACCCCGAGTGAATAATTAAGTACAGAAAGT 2401

Qy 2402 AGCTTGTTCAGAAATTTTTTGGATTTGGGAGCACTGTGTATCATACAGACATCTGT 2461

Db 2402 AGCTTGTTCAGAAATTTTTTGGATTTGGGAGCACTGTGTATCATACAGACATCTGT 2461

Qy 2462 TAGTGAGACACCAAAACCTGTGTACCCGTTTTCATGTATGAATTTTGTGTAGGT 2521

Db 2462 TAGTGAGACACCAAAACCTGTGTACCCGTTTTCATGTATGAATTTTGTGTAGGT 2521

Qy 2522 TGTCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAAGTGTATGAAAGGAGACATC 2581

Db 2522 TGTCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAAGTGTATGAAAGGAGACATC 2581

Qy 2582 TAAACAAATCCATTAGATTAACAGCTTCATGACGACAGAGGAAACTAATCTCAATGTT 2641

Db 2582 TAAACAAATCCATTAGATTAACAGCTTCATGACGAGAGGAAACTAATCTCAATGTT 2641

Qy 2642 TTAAGTAAATTAATCTGTACCTGTGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701

Db 2642 TTAAGTAAATTAATCTGTACCTGTGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701

Qy 2702 GGGGGGCGC 2710

Db 2702 GGGGGGCGC 2710

RESULT 4

189415

LOCUS 189415 2710 bp DNA linear PAT 10-AUG-1998

DEFINITION Sequence 8 from patent US 5721351.

ACCESSION 189415

VERSION 189415.1 GI:3409355

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2710)

AUTHORS Levinson,D,Adam.

TITLE Compositions and methods for the treatment and diagnosis of immune disorders

JOURNAL Patent: US 5721351-A 8 24-FEB-1998;

FEATURES

source 1..2710

BASE COUNT 789 a 592 c 644 g 678 t 7 others

ORIGIN

Query Match 99.9%; Score 2706.6; DB 6; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 242 TCAGAACTGATAAAGAAATGTGACATATCAAAATCCAGCAGATACAGCTTAAGGGCG 301

Qy 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGAAATGTGACTGTGATGACATGGGA 361

Db 302 ATCTCAACAAAGAGATGTGTCTGATCATTAAGAAATGTGACTGTGATGACATGGGA 361

Qy 362 CCTACTGCTGACAGATACAGTTCCTGGTCTTATGATATGAATTAATAATTAAGACTGAAT 421

Db 362 CCTACTGCTGACAGATACAGTTCCTGGTCTTATGATATGAATTAATAATTAAGACTGAAT 421

Qy 422 TAGACATCAAGACAGCAAGGTCATCCAGCTCAGACTGCGCATGGGAGCTGACTACAG 481

Db 422 TAGACATCAAGACAGCAAGGTCATCCAGCTCAGACTGCGCATGGGAGCTGACTACAG 481

Qy 482 CTCTCCAGAACCCCTAACACAGAGAGAAATGTTCAGAGACACAGACTGTGACCC 541

Db 482 CTCTCCAGAACCCCTAACACAGAGAGAAATGTTCAGAGACACAGACTGTGACCC 541

Qy 542 TCCATTAATACATGAAACAAATTTCCACATGGGCTGATGAATTAAGACTGTGAG 601

Db 542 TCCATTAATACATGAAACAAATTTCCACATGGGCTGATGAATTAAGACTGTGAG 601

Qy 602 AAACGATCAGAACTGCTATCCACATTTGAGTGGAGTCTCTGCTGGTTGACCTGGCAC 661

Db 602 AAACGATCAGAACTGCTATCCACATTTGAGTGGAGTCTCTGCTGGTTGACCTGGCAC 661

Qy 662 TTATCATTTGGTCTTAAATCCTTAATATGATATCTCTGTAAGAAAGAAATATGAGATT 721

Db 662 TTATCATTTGGTCTTAAATCCTTAATATGATATCTCTGTAAGAAAGAAATATGAGATT 721

Qy 722 TGAGCCTTATTAACCTGGCCAACTTGCTCCAGAGGGTGGCAATGACAGACATGCA 781

Db 722 TGAGCCTTATTAACCTGGCCAACTTGCTCCAGAGGGTGGCAATGACAGACATGCA 781

Qy 782 GGAATTCGCTCTGAGAAATATCTACACATGAGAGAGAACTATATGAAGTGAAGT 841

Db 782 GGAATTCGCTCTGAGAAATATCTACACATGAGAGAGAACTATATGAAGTGAAGT 841

Qy 842 CAAATGAGTACTACTGCTACAGTCAACAGCAGACAGCACTCTGACCTGACCTGCC 901

Db 842 CAAATGAGTACTACTGCTACAGTCAACAGCAGACAGCACTCTGACCTGACCTGCC 901

Qy 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGTGATTTCCCTTTGTGAAAACTATGTG 961

Db 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGTGATTTCCCTTTGTGAAAACTATGTG 961

Qy 962 ATATGCTACTTGGCAACCTCATTGGAGTTTGACCAAGCCACTGAGAAAAAGTTTCCA 1021

Db 962 ATATGCTACTTGGCAACCTCATTGGAGTTTGACCAAGCCACTGAGAAAAAGTTTCCA 1021

Qy 1022 GTTTTCTGGGATTAATTAATCAACAAAGGATTCATCTGATCACTGATCAATGAAAT 1081

Db 1022 GTTTTCTGGGATTAATTAATCAACAAAGGATTCATCTGATCACTGATCAATGAAAT 1081

Qy 1082 GCTCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCACATCCAGAGCTTCAATCAAG 1141

Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATTCGATCTCCACATCCAGAGCTTCAATCAAG 1141

Qy 1142 CGTGTGAAGCTCAGCTGCTGCTTCAATCATTAAGAAATGTGATGTGATGCTTTGAGA 1201

Db 1142 CGTGTGAAGCTCAGCTGCTGCTTCAATCATTAAGAAATGTGATGTGATGCTTTGAGA 1201

Qy 1202 CATAGAGTTTGTGATATCCGCAAGCTCTCTGAACAGTGAAGGGAATTAAGGGCTAA 1261

Db 1202 CATAGAGTTTGTGATATCCGCAAGCTCTCTGAACAGTGAAGGGAATTAAGGGCTAA 1261

Qy 1262 GATAGAGGTCGGCTTTTGTGATGTGAAAAATTTTAAAGAGTTGTGCTTTTCT 1321

Db 1262 GATAGAGGTCGGCTTTTGTGATGTGAAAAATTTTAAAGAGTTGTGCTTTTCT 1321

Qy 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGGGCAATGCTTAACAGATAT 1381

Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGGGCAATGCTTAACAGATAT 1381

[illegible]

QY	2462	TAGTGAGGACGACCAAACCTGTGGTACCGTTTTTTCATGTATGAAATTTGTGTGTTAGCT	2521
Db	2462	TAGTGAGGACGACCAAACCTGTGGTACCGTTTTTTCATGTATGAAATTTGTGTGTTAGCT	2521
QY	2522	TGCTTTAGCTAGCTAGCTGTGAGAGTCTCTGGCTTTCTTAAAGTGGTATGGAAGGAGACCATC	2581
Db	2522	TGCTTTAGCTAGCTAGCTGTGAGAGTCTCTGGCTTTCTTAAAGTGGTATGGAAGGAGACCATC	2581
QY	2582	TAACAAAATCATTAGAGATTAACAGCTCTCTCATGCAAGAGGAAACTTAATCTCAATGTT	2641
Db	2582	TAACAAAATCATTAGAGATTAACAGCTCTCTCATGCAAGAGGAAACTTAATCTCAATGTT	2641
QY	2642	TTAAAGTAATAAACTGACTGCGAAGAGTACTTTGACATATAAAAAAAAAAAAAAAAAAAG	2701
Db	2642	TTAAAGTAATAAACTGACTGCGAAGAGTACTTTGACATATAAAAAAAAAAAAAAAAAAAG	2701
QY	2702	GGCGGCCGC 2710	
Db	2702	GGCGGCCGC 2710	
RESULT 5	AF450241	2725 bp	mRNA linear ROD 11-FEB-2002
LOCUS	AF450241		
DEFINITION	Mus musculus T cell immunoglobulin mucin-3 mRNA, complete cds.		
ACCESSION	AF450241.1	GI:18182530	
VERSION			
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2125)		
AUTHORS	Monney, L., Sabatos, C.A., Gaglia, J.L., Ryu, A., Walchner, H., Chernova, T., Manning, S., Greenfield, E.A., Coyle, A.J., Sobel, R.A., Freeman, G.J. and Kuchroo, V.K.		
TITLE	Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease		
JOURNAL	Nature 415 (6871), 536-541 (2002)		
MEDLINE	11823861		
PUBMED	11823861		
REFERENCE	2 (bases 1 to 2725)		
AUTHORS	Monney, L., Sabatos, C., Gaglia, J.L., Ryu, A., Walchner, H., Chernova, T., Greenfield, E.A., Sobel, R.A., Freeman, G.J. and Kuchroo, V.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Adult Oncology, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115, USA		
FEATURES	Location/Qualifiers		
source	1..2725		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="AKR"		
	/db_xref="taxon:10090"		
	/cell_type="T cell"		
	67..512		
	/note="Tim-3"		
	/codon_start=1		
	/product="T cell immunoglobulin mucin-3"		
	/protein_id="AA165156.1"		
	/db_xref="GI:18182531"		
	/translation="MFGSLTLCNVALLIQLILARSLENGYKVEYKGNAYLPSCYTLPLPTLIDDTGYCCRIQPGELMADKLEIKLIDKAKVTPAQGTAEKSTTASPRITPTTRNGSLEDTLVTHNNNGTKISTWADKIDSGSETTATGAGVAGLTLALIGVLLIKMSSCKKKSSLSLITLAMPGLNAGAVRIRSENIYTLIEENVENSNBYCYANV		
	SQOPS"		
BASE COUNT	805 a 587 c 641 g 692 t		
ORIGIN			
Query Match	96.8%	Score 2624.6;	DB 10; Length 2725;
Best Local Similarity	99.0%;	Pred.No. 0;	



Matches 2655; Conservative 6; Mismatches 20; Indels 2; Gaps 2;	
QY	18 GGATTTCCCTCCCAAGTACTATGTTTTCAGGCTTTACCTCACTGTGTCTGTCT 77
Db	45 GGAATTTCCCTCCCAAGTACTATGTTTTCAGGCTTTACCTCACTGTGTCTGTCT 104
QY	78 GGTCAACTACTACTTGGCAAGGCTATTGGAAGTGTATTAAAGTTGAGTTGGTAA 137
Db	105 GCTGCAACTACTACTTGGCAAGGCTATTGGAAGTGTATTAAAGTTGAGTTGGTAA 164
QY	138 TGGCTATCTGCCCTGCAAGTTTACCTTACATCTGGGACATTTGTGCTATGTCTG 197
Db	165 TGGCTATCTGCCCTGCAAGTTTACCTTACATCTGGGACATTTGTGCTATGTCTG 224
QY	198 GGGCAAGGATTTGTCTTGTGTTCAAGTGTACCAATGAGTTCTCAAACTGATGA 257
Db	225 GGGCAAGGATTTGTCTTGTGTTCAAGTGTACCAATGAGTTCTCAAACTGATGA 284
QY	258 AATGTGACATATCAGAAATCCAGAGATPACAGCTAAAGGGGATCTCAACAAGAG 317
Db	285 AATGTGACATATCAGAAATCCAGAGATPACAGCTAAAGGGGATCTCAACAAGAG 344
QY	318 TGTGTCTGATCATTAAGAAATGTGACTGTGATGACCAATGGGACCTACTGTGAG 377
Db	345 TGTGTCTGATCATTAAGAAATGTGACTGTGATGACCAATGGGACCTACTGTGAG 404
QY	378 AAGTTCCCTGGCTTATGATTAATTAATAATTAAGTGAATTTGAATCAACAGAC 437
Db	405 AAGTTCCCTGGCTTATGATTAATTAATAATTAAGTGAATTTGAATCAACAGAC 464
QY	438 CAAGTCACTCCAGCTCAGACTGCCATGGGGACTCTACTACAGCTTCTCAAGAAC 497
Db	465 CAAGTCACTCCAGCTCAGACTGCCATGGGGACTCTACTACAGCTTCTCAAGAAC 524
QY	498 AACCGAGAGAGAAATGTTTCAAGACACAGACACTGTGACCTCCATTAATAAGT 557
Db	525 AACCGAGAGAGAAATGTTTCAAGACACAGACACTGTGACCTCCATTAATAAGT 584
QY	558 AACCAAAATTTTCCACATGGGCTGATGAATTAAGGACTCTGGGAAACGATCAAG 617
Db	585 AACCAAAATTTTCCACATGGGCTGATGAATTAAGGACTCTGGGAAACGATCAAG 644
QY	618 TATCCACATTTGAGTGGAGTCTGTGGGTGACCTCTGAGCACTTATCATTTGTT 677
Db	645 TATCCACATTTGAGTGGAGTCTGTGGGTGACCTCTGAGCACTTATCATTTGTT 704
QY	678 AATCTTTAAATGATTTCTGTGAAGAAAGAAATTAAGTGAAGTTGAGCTTATTA 737
Db	705 AATCTTTAAATGATTTCTGTGAAGAAAGAAATTAAGTGAAGTTGAGCTTATTA 764
QY	738 GGGCAACTGGCTCCAGAGAGGTTGGCAATGAGAGAGCATGAGATTCGCTTGA 797
Db	765 GGGCAACTGGCTCCAGAGAGGTTGGCAATGAGAGAGCATGAGATTCGCTTGA 824
QY	798 AATATCTACACCATGAGAGAGTATATGAAGTGAAGTCAATCAATGAGTACTG 857
Db	825 AATATCTACACCATGAGAGAGTATATGAAGTGAAGTCAATCAATGAGTACTG 884
QY	858 CTACGTCAACAGCCAGCAGCATCTGACCGCTCTGAGCTGCACTTTTAAAGCTCC 917
Db	885 CTACGTCAACAGCCAGCAGCATCTGACCGCTCTGAGCTGCACTTTTAAAGCTCC 944
QY	918 CTTCATTTTGACTTTGGTATTTCTTGTGGAAGAACTATGATGATGCACTTGGCA 977
Db	945 CTTCATTTTGACTTTGGTATTTCTTGTGGAAGAACTATGATGATGCACTTGGCA 1003
QY	978 CCTCATTTGAGGTTTGCACACAGCAGTGAAGAAAGATTCAGTTTTCTGGGAAAT 1037
Db	1004 CCTCATTTGAGGTTTGCACACAGCAGTGAAGAAAGATTCAGTTTTCTGGGAAAT 1063
QY	1038 TAACTCAAGGGGATTCGATGTGAATCTCATGCTAATGAAATGCTCAATTTATCC 1097
Db	1064 TAACTCAAGGGGATTCGATGTGAATCTCATGCTAATGAAATGCTCAATTTATCC 1123

QY	1098 GAGTTTCAGGATTCGATTCCTCACTCCAGAGACTTCAATCATTCGCTGTGAAGCTCACT 1157
Db	1124 GAGTTTCAGGATTCGATTCCTCACTCCAGAGACTTCAATCATTCGCTGTGAAGCTCACT 1183
QY	1158 CGTGTCTTATCATTAAGAAATGTTAGTGTATGTCTTGTGACATTAAGGTTGTGT 1217
Db	1184 CGTGTCTTATCATTAAGAAATGTTAGTGTATGTCTTGTGACATTAAGGTTGTGT 1243
QY	1218 ATATCCGCAAGCTCTGAAACAGGTAGGGGAAATAAGGCTTAAGATGAAGGTTGCG 1277
Db	1244 ATATCCGCAAGCTCTGAAACAGGTAGGGGAAATAAGGCTTAAGATGAAGGTTGCG 1303
QY	1278 CTTTGTGATGTTGGAATAATCTTAAAGAGTTGTAAGCTTTCTGAGATTTCTGAC 1337
Db	1304 CTTTGTGATGTTGGAATAATCTTAAAGAGTTGTAAGCTTTCTGAGATTTCTGAC 1362
QY	1338 GAAAGTTAAGAAAGCAAGGTGGCATATGCTTAAACAGATATTAAGTTGGAACTTGA 1397
Db	1363 GAAAGTTAAGAAAGCAAGGTGGCATATGCTTAAACAGATATTAAGTTGGAACTTGA 1422
QY	1398 GCAGAGGCTGATTAAGTTTCAAGGTCAAGCCAGGCTATGCTGTAAAGCTGTCAATC 1457
Db	1423 GCAGAGGCTGATTAAGTTTCAAGGTCAAGCCAGGCTATGCTGTAAAGCTGTCTCAAAATC 1482
QY	1458 CAAAGACCAAAATTAACATAGAGACAGCAGAGGCTGAGATGAGGCTTGGACATGAG 1517
Db	1483 CAAAGACCAAAATTAACATAGAGACAGCAGAGGCTGAGATGAGGCTTGGACATGAG 1542
QY	1518 TGCATTTGTATCAACACAGAGAACTATATTTGATCTGAGACCCCACTGAAAAAGCTA 1577
Db	1543 TGCATTTGTATCAACACAGAGAACTATATTTGATCTGAGACCCCACTGAAAAAGCTA 1602
QY	1578 GGCCTGTGTAGACATGCTTTGTAGACTCAAGAGATGGAAGTTAAAGCAACAGATCCC 1637
Db	1603 GGCCTGTGTAGACATGCTTTGTAGACTCAAGAGATGGAAGTTAAAGCAACAGATCCC 1662
QY	1638 CGGGGCTTGGCTGACATGCTTAAGCTTAAGTGTGATGCTCAAGTCCACAGAGCTCT 1697
Db	1663 CGGGGCTTGGCTGACATGCTTAAGCTTAAGTGTGATGCTCAAGTCCACAGAGCTCT 1722
QY	1698 GTCTCAAGTAAAGATGAGTGTGATCTGAGTCTGAGCCATGCTCAATGGGGTGTCTCTCT 1757
Db	1723 GTCTCAAGTAAAGATGAGTGTGATCTGAGTCTGAGCCATGCTCAATGGGGTGTCTCTCT 1782
QY	1758 CAGAAAGACATGACATGCTGCTGCAACACACACACACACACACACACACACAC 1817
Db	1783 CAGAAAGACATGACATGCTGCTGCAACACACACACACACACACACACACACAC 1842
QY	1818 ACACACACACATGAAATGAAAGTTCTCTGAGCTGCTCACTCTATTAACATGAT 187
Db	1843 ACACACACACATGAAATGAAAGTTCTCTGAGCTGCTCACTCTATTAACATGAT 1902
QY	1878 CTCTACAGACTCTCTCTGCTCTGTGAAGCATGAGTGGAGAGTGAAGAGCTCC 1937
Db	1903 CTCTACAGACTCTCTCTGCTCTGTGAAGCATGAGTGGAGAGTGAAGAGCTCC 1962
QY	1938 AGTAATTTATTCAGACATCAGAGGCTGAGCAGAAAGCTGGAAGTTCAGAGACATG 1997
Db	1963 AGTAATTTATTCAGACATCAGAGGCTGAGCAGAAAGCTGGAAGTTCAGAGACATG 2022
QY	1998 TGGCCAAACATGCGCAGACTCTTTTCAACAAGAAAGGTTTACCGCAAGAGCTGCTG 2057
Db	2023 TGGCCAAACATGCGCAGACTCTTTTCAACAAGAAAGGTTTACCGCAAGAGCTGCTG 2082
QY	2058 TCTGTAAAGGAAACCTTGCAGAAAGCAAACTTTGACTGTGTGTCTCAAGGGGAACTG 2117
Db	2083 TCTGTAAAGGAAACCTTGCAGAAAGCAAACTTTGACTGTGTGTCTCAAGGGGAACTG 2142
QY	2118 ACTCAGACACTTCTCATTTCTGTGAGGAACTGAGCTGTTTCTGACAGAAACACAC 2177
Db	2143 ACTCAGACACTTCTCATTTCTGTGAGGAACTGAGCTGTTTCTGACAGAAACACAC 2202

QY 2178 GGTGACTGGGACATACGAGGAGAGCTCTTGACGCAATCTATATATAGTCAGCAAAATATT 2237  
 Db 2203 GGTGACTGGGACATACGAGGAGAGCTCTTGACGCAATCTATATATAGTCAGCAAAATATT 2262  
 QY 2228 CTTTGGGAGGACAGTCGTCCACCAATTTGATTTCCAGCCGGTGGACCTTCAGTTTATCTG 2297  
 Db 2263 CTTTGGGAGGACAGTCGTCCACCAATTTGATTTCCAGCCGGTGGACCTTCAGTTTATCTG 2322  
 QY 2298 GCTTACAGCTGCTGCTCCAGTGCCTTTGATCTGTCTGGCTCCATCTATAACAGAAATCA 2357  
 Db 2323 GCTTACAGCTGCTGCTCCAGTGCCTTTGATCTGTCTGGCTCCATCTATAACAGAAATCA 2382  
 QY 2358 AATTAATATAGACCCGAGTGAATAATTAAGTGAAGACAGAAAGTAGCTTTGTTCAAGAT 2417  
 Db 2383 AATTAATATAGACCCGAGTGAATAATTAAGTGAAGACAGAAAGTAGCTTTGTTCAAGAT 2442  
 QY 2418 TTTTTCGATTTGGGAGACATCTGTATCATCAGAGACATCTGTTAGTGAAGACACCAA 2477  
 Db 2443 TTTTTCGATTTGGGAGACATCTGTATCATCAGAGACATCTGTTAGTGAAGACACCAA 2502  
 QY 2478 ACCTGTGATCCGCTTTTTCATGATGAATTTGTTTGGTGGCTTCTAGCTAGCTG 2537  
 Db 2503 ACCTGTGATCCGCTTTTTCATGATGAATTTGTTTGGTGGCTTCTAGCTAGCTG 2562  
 QY 2538 TGAAGTCTGCTGCTTTCTTAAAGTGGTATGAGAGGAGACCATCTAAGAAATCCATTAG 2597  
 Db 2563 TGAAGTCTGCTGCTTTCTTAAAGTGGTATGAGAGGAGACCATCTAAGAAATCCATTAG 2622  
 QY 2598 AGATACAGCTCTCATGAGAGAGGAAACTATCTCAATGCTTTTAAAGTAAATAACT 2657  
 Db 2623 AGATACAGCTCTCATGAGAGAGGAAACTATCTCAATGCTTTTAAAGTAAATAACT 2682  
 QY 2658 GTACGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAAAA 2700  
 Db 2683 GTACGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAAAA 2725

RESULT 6  
 AL669948  
 LOCUS  
 DEFINITION Mouse DNA sequence from clone RP23-248K2 on chromosome 11, complete sequence.  
 ACCESSION AL669948  
 VERSION AL669948.7 GI:19772854  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Tracey A.  
 Direct Submission  
 Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbrey@sanger.ac.uk; Clone requests: clonerequests@sanger.ac.uk  
 On Mar 27, 2002 this sequence version replaced gi:19521317.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP. Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep/RP23-248K2](http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP23-248K2) is from the RPI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6.  
 Location/Qualifiers  
 1..223030  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-248K2"  
 /clone\_11b="RPI-23"  
 BASE COUNT 63014 a 47122 c 47217 g 65677 t  
 ORIGIN

Query Match 69.6%; Score 1886.6; DB 10; Length 223030;  
 Best Local Similarity 97.7%; Pred. No. 0;  
 Matches 1941; Conservative 6; Mismatches 30; Indels 10; Gaps 3;  
 ORIGIN  
 QY 719 GTTGAAGCTTATTAACATGCGCAACTGCTCCAGAGGGTGGCAATGACAGAGAG 778  
 Db 55170 GTTGAAGCTTATTAACATGCGCAACTGCTCCAGAGGGTGGCAATGACAGAGAG 55229  
 QY 779 TCAGGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 838  
 Db 55220 TCAGGATTCGCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGAAGA 55289  
 QY 839 ATTCAATGAGTACTACTGTACGTCAACAGCCAGACCAATCTGACCGCTCTGACT 898  
 Db 55290 ATTCAATGAGTACTACTGTACGTCAACAGCCAGACCAATCTGACCGCTCTGACT 55349  
 QY 899 GCCACTTTTAAAGCTGCGCTCAATTTCTGACTTTGATTTCCCTTTTGTGAAAAATAT 958  
 Db 55350 GCCACTTTTAAAGCTGCGCTCAATTTCTGACTTTGATTTCCCTTTTGTGAAAAATAT 55408  
 QY 959 GTGATATGCTCACTTGGCAACCTCAATTGAGAGTTCGACACAGCACTGAGAAAAAGATT 1018  
 Db 55409 GTGATATGCTCACTTGGCAACCTCAATTGAGAGTTCGACACAGCACTGAGAAAAAGATT 55468  
 QY 1019 CCAAGTTTCTGGGATTAATTAATCTCAAGAGGATTCGACTGTAACTCATGCTACATTGA 1078  
 Db 55469 CCAAGTTTCTGGGATTAATTAATCTCAAGAGGATTCGACTGTAACTCATGCTACATTGA 55528  
 QY 1079 AATGCTCATTTTATCCCTGAGTTTCAGGAGTTCGATTCCTCCACCTCCAGAACCTTCAATC 1138  
 Db 55529 AATGCTCATTTTATCCCTGAGTTTCAGGAGTTCGATTCCTCCACCTCCAGAACCTTCAATC 555  
 QY 1139 ATGCGTGTGAAGCTCACTGCTGCTTTCATCATTAGGAATGTTAGTGTGCTTTG 1198  
 Db 55589 ATGCGTGTGAAGCTCACTGCTGCTTTCATCATTAGGAATGTTAGTGTGCTTTG 55648  
 QY 1199 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGGAACAGTGAAGGGAATAAAGGC 1258  
 Db 55649 AGACATAGAGGTTTGTGTATATCCGCAAGCTCTGGAACAGTGAAGGGAATAAAGGC 55708  
 QY 1259 TAAATAGAGAGGTCGCTTTTGTGTATGTTGAAAAATCTTTAAGAAGTGTGCTTT 1318  
 Db 55709 TAAATAGAGAGGTCGCTTTTGTGTATGTTGAAAAATCTTTAAGAAGTGTGCTTT 55767  
 QY 1319 TCTAGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGTGGCATATGCTTAACAGCA 1378  
 Db 55768 TCTAGAGATTTCTGACCTTGAAGATTAAGAAAAAGCAGGTGGCATATGCTTAACAGCA 55827  
 QY 1379 TATTAATTGGGACCTTAGGAGAGAGGTGATTAAGTGAAGTGAAGCCAGGCTATGCTG 1438  
 Db 55828 TATTAATTGGGACCTTAGGAGAGAGGTGATTAAGTGAAGTGAAGCCAGGCTATGCTG 55887  
 QY 1439 GTAAGACTGTCTCAACATCCCAAGCAAAATTAACATATAGACACAGAGGCTGAGCA 1498  
 Db 55888 GTAAGACTGTCTCAACATCCCAAGCAAAATTAACATATAGACACAGAGGCTGAGCA 55947  
 QY 1499 TGAGGCTCGACAGTGAAGTGCATTTGTGTATACAGACAGGAATCTATATTTGATGCTAG 1558

Db 55948 TGAAGCTCGGACAGTGGATGATTTTGTACAAAGCAGAGAAATCTATATTTGATGTAG 56007  
 Oy 1559 ACCCCACATGAAAAAGCTAGGCTGTGTAGAGCATGCTTTAGACTCAAGAGATGAGAG 1618  
 Db 56008 ACCCCACATGAAAAAGCTAGGCTGTGTAGAGCATGCTTTAGACTCAAGAGATGAGAG 56067  
 Oy 1619 TAAAGGACACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGAGTTC 1678  
 Db 56068 TAAAGGACACAGATCCCGGGGGCTTGGGTGACATGCTTACCTAGGTGCTGAGTTC 56127  
 Oy 1679 CAAGTCCACAAAGAGTCCCTGTCTCAAGATGAGTGGATGATCTGGGACATGACAT 1738  
 Db 56128 CAAGTCCACAAAGAGTCCCTGTCTCAAGATGAGTGGATGATCTGGGACATGACAT 56187  
 Oy 1739 GGGGGTGTCTCTCTCTCAGAGAGACATGACATGACATGACATGACATGACATGACAT 1798  
 Db 56188 GGGGGTGTCTCTCTCTCAGAGAGACATGACATGACATGACATGACATGACATGACAT 56247  
 Oy 1799 CAC 1850  
 Db 56248 CAC 56307  
 Oy 1851 TGCTGTCT 1910  
 Db 56308 TGCTGTCT 56367  
 Oy 1911 ATGAGTGGAGACATGAC 1970  
 Db 56368 ATGAGTGGAGACATGAC 56427  
 Oy 1971 AGAAGGTGAGAGATTCAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2030  
 Db 56428 AGAAGGTGAGAGATTCAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56487  
 Oy 2031 AAAAGGTGAGAGATTCAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2090  
 Db 56488 AAAAGGTGAGAGATTCAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56547  
 Oy 2091 TGAAGTGTGTGTCTCAAGGGGAACTGACATGACACAACTTCTCTCTCTGAGAGAACT 2150  
 Db 56548 TGAAGTGTGTGTCTCAAGGGGAACTGACATGACACAACTTCTCTCTCTGAGAGAACT 56607  
 Oy 2151 GGAAGCTGTGTGTCTCAAGGGGAACTGACATGACACAACTTCTCTCTCTGAGAGAACT 2210  
 Db 56608 GGAAGCTGTGTGTCTCAAGGGGAACTGACATGACACAACTTCTCTCTCTGAGAGAACT 56667  
 Oy 2211 AGCAATCTATATAGTACAGCAAAATTTCTTTGGGAGACAGTGTGACCAAAATTTGATTC 2270  
 Db 56668 AGCAATCTATATAGTACAGCAAAATTTCTTTGGGAGACAGTGTGACCAAAATTTGATTC 56727  
 Oy 2271 CAAGCGGTGAGACCTGAGTTTCATCTGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTG 2330  
 Db 56728 CAAGCGGTGAGACCTGAGTTTCATCTGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTG 56787  
 Oy 2331 TGTGTGCTCCCATCTATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2390  
 Db 56788 TGTGTGCTCCCATCTATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 56847  
 Oy 2391 ACCAGAAAGTGTGCTTCAAGATTTTCTTTGCAATGGGAGCAACTGTGACATCAG 2450  
 Db 56848 ACCAGAAAGTGTGCTTCAAGATTTTCTTTGCAATGGGAGCAACTGTGACATCAG 56907  
 Oy 2451 AGGACATCTGTAGTACAGACACCAAAACCTGTGTACCGTTTCTCATGTATGAAATTT 2510  
 Db 56908 AGGACATCTGTAGTACAGACACCAAAACCTGTGTACCGTTTCTCATGTATGAAATTT 56967  
 Oy 2511 GTTGTGTAGTGTCTCTAGT 2570  
 Db 56968 GTTGTGTAGTGTCTCTAGT 57027  
 Oy 2571 GGGAGACATCTAACAAATCCATTGAGATTAAGCTCTCATGTACAGAGAGAACTTAA 2630

Db 57028 GGGAGACATCTAACAAATCCATTGAGATTAAGCTCTCATGTACAGAGAGAACTTAA 57087  
 Oy 2631 TCTCAAAATGTTTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2690  
 Db 57088 TCTCAAAATGTTTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 57147  
 Oy 2691 AAAAAA 2697  
 Db 57148 AGATCAA 57154  
 RESULT 7  
 AC119505/C  
 LOCUS  
 DEFINITION  
 Rattus norvegicus c1one CH230-344M17, WORKING DRAFT SEQUENCE, 3  
 unordereed pieces.  
 AC119505  
 AC119505.5 GI:25138135  
 HTG: HTGS\_PHASE1, HTGS\_DRAFT, HTGS\_FULFILLTOP.  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 177087)  
 Muzny,D,Marie., Metzker,M, Lee., Abrazon,S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsdorfs, S., Amth, A., Angilano, D.,  
 Anyalebechi, V., Ayogei, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Bissalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, R., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Becotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,  
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M.,  
 Hollins, B., Howells, S., Huliy, S., Hume, J., Ididit, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenshewe, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Maniyan, S., McLeod, M. P., McNelly, T. Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Mundiase, M., Murphy, M., Nat, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,  
 Nwakoelimen, O., Okumotu, G., Olarnunagsoon, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkocho, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,  
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. D.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisom, I., Sitter, C. D., Smjs, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K.,  
 Vais, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,  
 Williams, G., Willson, R., Wlezyk, R., Woodson, H., Woley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,





Db	65474	GTGCATCCTTGTGAACCTCCAGAGATAGAGAGGTAAAGGCAACACATCCCTGGGCATG	65415
OY	16447	CGTGCAGTCAAGCTTAGCCTTAGGTGCTGAATTCAGATGCCAAGAGTCCCTGTCTCAGAG	1706
Db	65414	ATTGACCCCACTTTAGCCTTAGATGGTGAGTTGACGAATCCGTGAGATCCCTGTCTCATATA	65355
OY	1707	TAAGATGGATCTGATCTGTGGCGATATCCATATGGGGTGTGCTCTCTCTCAGAAAG----	1763
Db	65354	TAAATGGATCTGTCTCTGGGGAACATCAATGAGACGTCTCTGGGACTCTCTGCAGACA	65295
OY	1764	-----AGACATGCACATGCCCCCTGCACACACACACACACACACACACAC	1809
Db	65294	TGCGCATGTACTTACACACGACACATCGCATCGCACGCAATGGCGACGCGCACATGA	65235
OY	1810	ACACACACACACACACACATGMA-----TGAGG	1840
Db	65234	GGTTACACGCAAAAGATTAAAGAAAGAAAGATTCAGAACAGGGCTAAATGATATGATA	65175
OY	1841	GTTTCTCTCTGTGCTGTACTCTCTCTATACATATATCTTACAGACATCTCTGTGCTCT	1900
Db	65174	GTTCTCTGTGTGCTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	65115
OY	1901	CTGTTAAGACA-----TGAGTGGAGCATGAGCAGAGCAGTCCAGTAAATTAATCCAGCAC	1955
Db	65114	CTGTTAAGACCTGCTCTGAAAGGAGACATGAAATGACACCTGTAAATTAATTCAGCAC	65055
OY	1956	TCAGAAAGCTTGAGCAGAAAGCGTGGAGAGTTTCAGAGACATGTGCCCAACATCTGCCAGAC	2015
Db	65054	CCAGAAAGTTTAAACAGAGAGATGAGAGAGGTTCAGAGACATGTGTCTTACATCTGCCAGAC	64995
OY	2016	TCTTCTTACACAAACAAAAGTTTACCCGCAAGACCTGTGTCTGTAAAGAAAC--C	2073
Db	64994	TC-TCGTACACAAACAAAAGTTTACCCGCAAGACCTGTGTCTGTAAAGAAACCTTT	64936
OY	2074	CTGGAAGGCAAACTTTGACTGTTGTGTGCTCAAGGGGAACCTGACCTCAGACAACTTCTC	2133
Db	64935	TTACGAAAGCTGCACCTTGTCTGTGTGTCTCAAGGGGAACCTGACCTCAGACAACTTCTC	64876
OY	2134	CATTCTTGAGAAACTGTGAGC--TGTTTTGACAGAAAGAACAAACCGGTGACTGGAGCAT	2191
Db	64875	CATTCTTGAGAAAGAACTGTGAGCCTGTGTTCTTGCACAGAAAGAAACCGGTGACTGGAGCAT	64816
OY	2192	ACGAAAGCAGAGCTCTTGACAGCATCTATATAGTCAGCAAAATATTTCTTTGGAGAGACA	2250
Db	64815	ACGAAAGCAGAGCTCTTGACAGCATCTATATAGTCAGCAAAATATTTCTTTGGATAGACA	64756
OY	2251	GTCGCACCAAA--TTGATTTCCAAAGCCGGGAGCCTGAGTTGATCTGCGCTTACAGCTGC	2309
Db	64755	GTCGCACCAAACTTAATTTCCAAAGCTGGAGAGCTGAGTTGATCTGCGCTTACAGCTGC	64696
OY	2310	CTGCCAGTGCCCTTGATCTGTGTGCTGCTCCCATCTTAAACAGATCAAAATTAATATAGAC	2369
Db	64695	TTGGCAGATG--CCTTGGCTGTGTGTGTTCTCATATATACAGATCT--AATTAATATAGAC	64638
OY	2370	CCCGATGTGAATATTTAAATGAGACAGAAAGTGAAGCTTTGTTCAAAGATTTTGTGATTTG	2429
Db	64637	CCAGATGTGAATATTTAAATGAGACAGAAAGTGAAGCTTTGTTGAAAGC--CTTTTGCACCG	64579
OY	2430	GGGAGCAACGTGTATCATCAGAGACATCTGTTATAGTGAAGACACCAAAACCTG--TGATTA	2487
Db	64578	GGGAGCAACGTGTATGTCAAGAGACGTCTGTTATAGTGAAGACCTGAACCTGCAATGTTT	64519
OY	2488	CCGTTTTTTTCAATGATGATTTTGTGTTTAAAGTGTCTTACACTAGCTGTGAGAGTCTCT	2547
Db	64518	GGATCTTTTCTTGTATGATTTTGTATGTTTAAAGCTGATTTTACACTAGCTGTGAGACTACT	64455
OY	2548	GGCTTTTCTTAAAGTGGATATGAGAGGAGACCATCTTAAACAAATCTATAGATTAACAC	2607
Db	64458	GGATCTTTTAAAGTGGATATGAGAGGAGACCATCTGCAATCAATT--AGTTAATCATC	64402
OY	2608	TCTCATGCAAGAGGAAACCTATCTCAATGTTTTTAAATTAATA--AATCTGACTG	2663

Db	64401	TCTCAGTCAGCATGGGAAAATTAATCACTGAGTGTTTAAAGTAATAAACAACATGTA	64342
Oy	2664	GCAAACTACTTTAGCATATAAAAA	2697
Db	64341	GTGAAGTACTTTGAGCATATTTAAACAGATCAA	64308
RESULT 8			
AC098539/c			
LOCUS			
DEFINITION			
AC098539	233370 bp	DNA	linear
AC098539/c			
AC098539			
AC098539	GI:30521658		
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.			
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.			
Rattus norvegicus			
Rattus norvegicus			
Rattus norvegicus			
Enkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE			
1 (bases 1 to 233370)			
AUTHORS			
Muzny,D,Marle, Metzker,M, Lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D, Arayaalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benamed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas-V, Carter,K, Cavazos,I, Cessar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,N, Hernandez,J, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howells,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levam,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensueta,L, Loulsegue,H, Lozado,R,J, Lu,X, Ma,J, Meshawati,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E, Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,B, Mlosoavljetic,A, Miner,G, Mijta,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundaasa,M, Murphy,M, Nait,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nsoekilemeh,O, Okwuonu,G, Olarnpusongon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poindecker,A, Popovic,D, Prims,E, Pu,L, Plazo,M, Quiroz,O, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Rutz,S,J, Sanders,W, Savary,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartbeyn,A, Sisson,I, Sitter,C,D, Smaj,D, Sneed,A, Sodergren,E, Song,X-Z, Sorrelle,R, Soza,D, Steinle,M, Strong,R, Sulton,A, Svatek,A, Tabori,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tinley,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasean,D, Waldron,L, Walker,B, Wang,J, Wang,O, Wang,S, Warren,R, Warren,R, Wei,X, White,F, Williams,G, Willison,R, Wlczek,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,D, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.			
TITLE			
JOURNAL			
REFERENCE			
2 (bases 1 to 233370)			

```

AUTHORS      Worley,K.C.
TITLE        Direct Submission
JOURNAL      Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
REFERENCE    3 (bases 1 to 233370)
AUTHORS      Rat Genome Sequencing Consortium.
TITLE        Direct Submission
JOURNAL      Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
              On May 10, 2003 this sequence version replaced gi:23268780.
              The sequence in this assembly is a combination of BAC based reads
              and whole genome shotgun sequencing reads assembled using Atlas
              (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
              in the feature table below represents a scaffold in the Atlas
              assembly (a 'contig-scaffold'). Within each contig-scaffold,
              individual sequence contigs are ordered and oriented, and separated
              by sized gaps filled with Ns to the estimated size. The sequence
              may extend beyond the ends of the clone and there may be sequence
              contigs within a contig-scaffold that consist entirely of whole
              genome shotgun sequence reads. Both end sequences and whole genome
              shotgun sequence only contigs will be indicated in the feature
              table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: CH1Y
Center clone name: CH230-94D16

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 195507 bases at least Q40
Consensus quality: 199408 bases at least Q30
Consensus quality: 202062 bases at least Q20
Estimated insert size: 199541; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 208325: contig of 208325 bp in length
* 208326 208425: gap of unknown length
* 208426 230463: contig of 22038 bp in length
* 230464 230563: gap of unknown length
* 230564 231796: contig of 1233 bp in length
* 231797 231896: gap of unknown length
* 231897 233370: contig of 1474 bp in length.
*
* Location/Qualifiers
* 1. 233370
  /organism="Rattus norvegicus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10116"
  /clone="CH230-94D16"
  1. 1090
  /note="wgs contig"
  misc_feature 33723..35092
  /note="wgs_contig"
  misc_feature 48213..50325
  /note="wgs_contig"
  misc_feature 62570..63741
  /note="wgs_contig"
  111215..112292
  /note="wgs_contig"

```

	misc_feature	116474..117898	/note="wgs_contig"	
	misc_feature	22897..23043	/note="wgs_contig"	
BASE COUNT	58498 a	44016 c	43600 g	57781 t 29475 others
ORIGIN				
Query Match	42.2%	Score 1144.8;	DB 2; Length 233370;	
Best Local Similarity	79.8%;	Pred. No. 0;		
Matches 1656;	Conservative 4;	Mismatches 326;	Indels 88;	Gaps 23;
OY	690	GTATTCCTGTAAGAAAAAAGATTACAGTTTGAGCCCTTATTACATCGCAACTGGCC	749	
Db	192212	GCTCTTCGCCAACCTTATCATTATTTTTTACAGTCTTATTAACAATGGCACTCCCC	192213	
OY	750	TCCAGAGAGGTGGCAAAATGCAGAGCAGTCAAGATTGCTCTGAGAAAATATCTTAC	809	
Db	192212	ACCAGAGAGGTGGGAATGCAGAGCAGGAGCATTCGGTCTGAGAAAATCATCTAC	192213	
OY	810	CATGAGAGAGAAGCTATATGAAATGAGAAATTTCAAATGAGTCTACTGACTCACAG	869	
Db	192152	TATGAGAGAGAATATATGAAATGAGAAATTTCAAATGAGTCTACTGACTCACAG	192093	
OY	870	CCAGAGCAATCTGACCGCCCTGGATGTCGCACCTTTAAAGGCTCGCCTTCAATTCTGA	929	
Db	192092	CCAGAGCAATCTGACCGCCCTGGGCTCCCGACTTTAAAGGCTCCCTTCAATTCTGA	192033	
OY	930	CTTTGTAATTCCTTTTGTGAAAACTATGTGATATGTCACTTGGCAACTCATTTGAGG	989	
Db	192032	CTTTGTAATTCCTTTTGTGAAAACTATGTGATATGTCACTTGGCAACTCATTTGAGG	191974	
OY	990	TTCTGACCAACCCAATGAGAAAGAGTTCCAGTTTCTGGGATTAATTAATCTCACAG	1049	
Db	191973	TTCTGACCAACCCAATGAGAAAGAGTTCTGTTTCTGGGATTAACCAATCTCACAG	191914	
OY	1050	GGATTTCAGCTGAATCTCAATGCTACATTGAATGCTCAATTTATCCCTGAGTTTCAAGG	1109	
Db	191913	AGATTTCAGCTGAATCTCAATGCTACATTGAATGCTCAATTTATCCCTGAGTTTCAAGG	191856	
OY	1110	TCGGAATCTCCCACTCCAGAGACTTCAATCAATGCTGTTGAAGCTCACTGTCCTTTCATA	1169	
Db	191855	CAGGAATCTCCCAACCCAGAGACTCAATCAATGCTGTTGAAGCTCACTGTCCTTTCATA	191796	
OY	1170	CATTAGGAATGTTATGATGTCCTTTGAGACATAGAGTTTGTGTTATATCCGAAG	1229	
Db	191795	CATTAGGAATGTTATGATGTCCTTTGAGACATTAAGTTTGTGTTATATCCGAAG	191729	
OY	1230	CTCCGGAACAGTAGGGGGAATTAAGGCTAAGATAGGAAGGTGC--GGCTTTGTTGAG	1288	
Db	191735	CTCCGGAACAGTAGGGGGAATTAAGGCTAAGATCGTAGGTGACGGCTTTGTTGAG	191676	
OY	1289	TTGGAATATCTTAAAGAGTTGTTGCTTTTCTAGAGATTTTGAACCTTGAAGATTAAG	1348	
Db	191675	CTCGAAA--TCTAATAAAGCTGTGTGTTTCTTABAAGTTCTTAACCTTGAAGATTAAG	191618	
OY	1349	AAAAAGCCAGGTGCGATATGCTTTAAACAGCATATAAATTGGGAACCTTAGCGAGAGGTG	1408	
Db	191617	----AGCGGGGTGCGATATGCTTTAAATTCATTACTTGGGAACCTTAGCGAGAGGTG	191522	
OY	1409	ATAAGTTTAAAGTCAAGCCAGGGCTTATGCTGTGTAAC--TGTCTCAMCATCCAAAGACA	1466	
Db	191561	GTGAGTTTAAAGTCAAGCTGAGGGCTTATGTTGTGTAACCTTTGTCTCAAAAACAAAGAGA	191502	
OY	1467	AAATTAACATAGAGCAGCAGAGGCTGAGATGAGGCTCGACAGTGTGATTTGTG	1526	
Db	191501	AAATTAACATAGAGCA--CTGAGACTGAGATATGCT---CACTGAGGTGC--TTGTT	191448	
OY	1527	TACAAGCAGAGAACTTATATTGATGCTAGACCCCATGAATAAGCTAGAGCTTGTA	1586	
Db	191447	ATTAAGCATGAGAACTGAGTTTGTCTCCAGACTGACATTAATAAGCTAGAGCTTGTA	191388	
OY	1587	GAGCATGCTTTGAGCTCAAGAGATGAGAGGTTAAAGCAACAGATCCCGGGGCTTG	1646	

```

Db 191387 GTGCATGCTTGTGAACTCCAGAGATAGGAGGTAAAGGCAAGACAGATCCCTGGGGCATG 191328
Qy 1647 CGTGCACTGAGCTTACCTTACCTGAGTGTGCTCAAGTCCCAAGAGTCCCTGTCTCAAG 1706
Db 191327 ATTGAGCCACCTTACCTTACCTGAGTGTGCTCAAGTCCCTGTCTCAAG 191268
Qy 1707 TAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1763
Db 191267 TAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 191208
Qy 1764 -----AGACATGCAATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1809
Db 191207 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 191148
Qy 1810 ACACACACACACACACACACATGAA-----TGAG 1840
Db 191147 GGTAAACCCCAAAAGATTAAGAAAGAGATCAGAAACGGGCTAAATGATAGTAA 191088
Qy 1841 GTTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1900
Db 191087 GTTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191028
Qy 1901 CTGTTAAGACA-----TGAGTGGAGCATGAGCAGTCCAGTAATTTATTCAGCAC 1955
Db 191027 CTGTTAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190968
Qy 1956 TCGAAGGCTGAGCAGAGAGCTGAGAGTTCAGAGCACTGCTCCACACCTGCTGCTGCTGCTGCT 2015
Db 190967 CCGAAGGCTGAGCAGAGAGCTGAGAGTTCAGAGCACTGCTCCACACCTGCTGCTGCTGCTGCT 190908
Qy 2016 TCTTCTTACACAGAAAGAGTACCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
Db 190907 TC-TGCTACACAGAAAGAGTACCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190849
Qy 2074 CTGCGAAGGCAAACTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2133
Db 190848 TTAAGAACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190789
Qy 2134 CATTCCTGAGAGAACTGAGC--TGTTCTGACAGAAACACCGGTGACCTGGAGCAT 2191
Db 190788 CATTCCTGAGAGAACTGAGC--TGTTCTGACAGAAACACCGGTGACCTGGAGCAT 190729
Qy 2192 AGAAGGCAAGAGCTTGTGAGCAATCTATATAGTACAGAAATATTC-TTGGGAGGACA 2250
Db 190728 AGAAGGCAAGAGCTTGTGAGCAATCTATATAGTACAGAAATATTC-TTGGGAGGACA 190669
Qy 2251 GTGCTGACCAAA-TTGATTTCCAGGCGGTGACCTGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2309
Db 190668 GTGCTGACCAAA-TTGATTTCCAGGCGGTGACCTGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190609
Qy 2310 CTGCCAGTGCCTTGAATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2369
Db 190608 CTGCCAGTGCCTTGAATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190551
Qy 2370 CCCGAGTGAATAATTAAGAGAGCAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429
Db 190550 CCCGAGTGAATAATTAAGAGAGCAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190492
Qy 2430 GGGAGCAACTGTGTACATCAGAGACATCTGTTAGTAGAGACACCAAACTG--TGTA 2487
Db 190491 GGGAGCAACTGTGTACATCAGAGACATCTGTTAGTAGAGACACCAAACTG--TGTA 190432
Qy 2488 CCGTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2547
Db 190431 CCGTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190372
Qy 2548 GCGTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2607
Db 190371 GCGTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190315
Qy 2608 TCTCATGACAGAAAGGAAATTAATCTCAATGTTTTTAAAGTAAAT--AACTGTACTG 2663
Db 190314 TCTCATGACAGATGAGAAATTAATCTCAATGTTTTTAAAGTAAATCAAACTGTACT 190255

```

```

Qy 2664 GCAAGTACTTTGACATTAATAAAAAAAAAAAAA 2697
Db 190254 GTGAAGTACTTTGACATTAATAAACAGATCA 190221

RESULT 9
AF399831
LOCUS AF399831 862 bp mRNA linear ROD 08-APR-2002
DEFINITION Mus musculus strain DBA/2 TIM3 mRNA, partial cds.
ACCESSION AF399831
VERSION AF399831.1 GI:17146680
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 862)
McIntire,J.J., Umetsu,S.E., Albani,O., Potter,M., Kuchroo,V.K.,
Barsh,G.S., Freeman,G.J., Umetsu,D.T. and Dekruyff,R.H.
Identification of Tgpr (an airway hyperreactivity regulatory locus)
and the linked Tim gene family
Nat. Immunol. 2 (12), 1109-1116 (2001)

JOURNAL
MEDLINE 21582130
PUBMED 11725301

REFERENCE
2 (bases 1 to 862)
McIntire,J.J., Umetsu,S.E., Albani,O., Potter,M., Barsh,G.S.,
Freeman,G.J., Umetsu,D.T. and Dekruyff,R.H.
Direct Submision
Submitted (12-JUN-2001) Pediatrics, Stanford University, 300
Pasteur Drive, Grant S-303, Stanford, CA 94305, USA
location/Qualifiers
1. .862
/organism="Mus musculus"
/mol_type="mRNA"
/strain="DBA/2"
/db_xref="taxon:10090"
/chromosome="11"
/map="between D1Mit271 and D1Mit22"
/tissue_type="spleen"
17. .862
/note="contains T-cell immunoglobulin domain and mucin
domain"
/codon_start=1
/product="TIM3"
/protein_id="AAL35776.1"
/db_xref="GI:17146681"
/translation="MPSGLTNCVLLQLLLARSLENAYVEGKNAYLPCSYLST
PGALVPMCKGKGFPMSCQTNELRTDRNTVYKRSRYLKGDLKGVSLIKNVT
LDDHGTVCRIOPFGIANDKLELKDIAKATPAQTAHDSPTASPTLTERNGE
ETQTLVTHNNNGTKISTWADBEIKDSGETIRTAIHIGVAGSLTALILGVILKRY
SCKRRKLSLSLITLANLPQGLAAGAVRIKSEENIYITSENYEVENSNBYCYN
SOQPS"

BASE COUNT 254 a 197 c 198 g 213 t
ORIGIN
Query Match 31.2%; Score 844.4; DB 10; Length 862;
Best Local Similarity 98.7%; Pred. No. 2,1e-233;
Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 1 CCCCTCCCAAGTACATGATTTTTCAGGTCTTACCTCAACGTGTGCTGCTGCTGCA 83
24 CCCCTCCCAAGTACATGATTTTTCAGGTCTTACCTCAACGTGTGCTGCTGCTGCA 60
1 CCCCTCCCAAGTACATGATTTTTCAGGTCTTACCTCAACGTGTGCTGCTGCTGCA 60
Qy 84 ACTACTACTTTCAGAGTCACTTGAAGATGTTTAAAGTGAAGTGTGTTAAATGCCCTA 143
Db 61 ACTACTACTTTCAGAGTCACTTGAAGATGTTTAAAGTGAAGTGTGTTAAATGCCCTA 120
Qy 144 TCTGCCCTGCACTTACCTTACATCTGAGGACACTTGTGCTGCTGCTGCTGCTGCA 203
Db 121 TCTGCCCTGCACTTACCTTACATCTGAGGACACTTGTGCTGCTGCTGCTGCTGCA 180
Qy 204 GGGATTCTGTCTTGTGTCACTGTACCAATGAGTTGCTCAGAACTGATGAAGAAATGT 263

```



TITLE Compositions and methods for the treatment and diagnosis of immune disorders

JOURNAL Patent: US 6288218-A 22 11-SEP-2001;

FEATURES Location/Qualifiers

source 1. 843

/organism="unknown"

BASE COUNT 253 a 185 c 196 g 209 t

ORIGIN

Query Match 31.1%; Score 843; DB 6; Length 843;

Best Local Similarity 100.0%; Pred. No. 5.3e-233;

Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ATGTTTTCAGGCTTACCTCTCAACCTGTCTCTGCTCTGCACTACTACTCTGCAAG 99

DB 1 ATGTTTTCAGGCTTACCTCTCAACCTGTCTCTGCTCTGCACTACTACTCTGCAAG 60

QY 100 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 159

DB 61 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 120

QY 160 ACTCTACCTACATCTGGGACCTTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 219

DB 121 ACTCTACCTACATCTGGGACCTTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 180

QY 220 TCACAGTGTACCAATGAGTTGCTCAAGACTGATGAAAGAAATGTGACATATTCGAATCC 279

DB 181 TCACAGTGTACCAATGAGTTGCTCAAGACTGATGAAAGAAATGTGACATATTCGAATCC 240

QY 280 AGCAGATACCAAGTGTAAAGGCGATCTCAACAAAGAGATGTCTCTGATCATTAAGAT 339

DB 241 AGCAGATACCAAGTGTAAAGGCGATCTCAACAAAGAGATGTCTCTGATCATTAAGAT 300

QY 340 GTGACTCTGATGATCCATGGGACCTTCTGCTGAGGATACATTTCCCTGCTTATGAAT 399

DB 301 GTGACTCTGATGATCCATGGGACCTTCTGCTGAGGATACATTTCCCTGCTTATGAAT 360

QY 400 GATTAATAATTAACTGAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 459

DB 361 GATTAATAATTAACTGAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 420

QY 460 GCCCATGGGACCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAGAAATGTTCA 519

DB 421 GCCCATGGGACCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAGAAATGTTCA 480

QY 520 GAGACACAGACACTGTGACCTTCCATTAATTAACAAATGAAACAAAATTTTCCATGGGCT 579

DB 481 GAGACACAGACACTGTGACCTTCCATTAATTAACAAATGAAACAAAATTTTCCATGGGCT 540

QY 580 GATGAATAATTAAAGATCTGGAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 639

DB 541 GATGAATAATTAAAGATCTGGAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 600

QY 640 TCTGCTGGGTTGACCTCTGCACTTATCATTTGTTGTTTAACTTAAATGTTTCTGT 699

DB 601 TCTGCTGGGTTGACCTCTGCACTTATCATTTGTTGTTTAACTTAAATGTTTCTGT 660

QY 700 AAGAAAAGAAAGTATGAGTTGAGCTTTATTAACCTGGCCAACTTGCCTCCAGAGAGG 759

DB 661 AAGAAAAGAAAGTATGAGTTGAGCTTTATTAACCTGGCCAACTTGCCTCCAGAGAGG 720

QY 760 TTGGCAAAATGCGAGACAGTCAAGATTCGCTGAGAGAAAATTCATACCAATCGAGAG 819

DB 721 TTGGCAAAATGCGAGACAGTCAAGATTCGCTGAGAGAAAATTCATACCAATCGAGAG 780

QY 820 AACGTATATGAGTGAAGATTAATCAATGAGTATTAATGCTGATCAACAGCAGAGCA 879

DB 781 AACGTATATGAGTGAAGATTAATCAATGAGTATTAATGCTGATCAACAGCAGAGCA 840

QY 880 TCC 882

DB 841 TCC 843

RESULT 12

AR232684

LOCUS AR232684 843 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 22 from patent US 6455685.

ACCESSION AR232684

VERSION AR232684.1 GI:27274961

KEYWORDS

SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 843)

AUTHORS Levinson, D.A.

TITLE Compositions and methods for the treatment and diagnosis of immune disorders

JOURNAL Patent: US 6455685-A 22 24-SEP-2002;

FEATURES Location/Qualifiers

source 1. 843

/organism="unknown"

BASE COUNT 253 a 185 c 196 g 209 t

ORIGIN

Query Match 31.1%; Score 843; DB 6; Length 843;

Best Local Similarity 100.0%; Pred. No. 5.3e-233;

Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ATGTTTTCAGGCTTACCTCTCAACCTGTCTCTGCTCTGCACTACTACTCTGCAAG 99

DB 1 ATGTTTTCAGGCTTACCTCTCAACCTGTCTCTGCTCTGCACTACTACTCTGCAAG 60

QY 100 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 159

DB 61 TCATTGGAAGATGTTATTAAGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 120

QY 160 ACTCTACCTACATCTGGGACCTTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 219

DB 121 ACTCTACCTACATCTGGGACCTTGTGCTATGTGCTGGGCAAGGATTTCTGCTTGG 180

QY 220 TCACAGTGTACCAATGAGTTGCTCAAGACTGATGAAAGAAATGTGACATATTCGAATCC 279

DB 181 TCACAGTGTACCAATGAGTTGCTCAAGACTGATGAAAGAAATGTGACATATTCGAATCC 240

QY 280 AGCAGATACCAAGTGTAAAGGCGATCTCAACAAAGAGATGTCTCTGATCATTAAGAT 339

DB 241 AGCAGATACCAAGTGTAAAGGCGATCTCAACAAAGAGATGTCTCTGATCATTAAGAT 300

QY 340 GTGACTCTGATGATCCATGGGACCTTCTGCTGAGGATACATTTCCCTGCTTATGAAT 399

DB 301 GTGACTCTGATGATCCATGGGACCTTCTGCTGAGGATACATTTCCCTGCTTATGAAT 360

QY 400 GATTAATAATTAACTGAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 459

DB 361 GATTAATAATTAACTGAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 420

QY 460 GCCCATGGGACCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAGAAATGTTCA 519

DB 421 GCCCATGGGACCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAGAAATGTTCA 480

QY 520 GAGACACAGACACTGTGACCTTCCATTAATTAACAAATGAAACAAAATTTTCCATGGGCT 579

DB 481 GAGACACAGACACTGTGACCTTCCATTAATTAACAAATGAAACAAAATTTTCCATGGGCT 540

QY 580 GATGAATAATTAAAGATCTGGAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 639

DB 541 GATGAATAATTAAAGATCTGGAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAG 600

QY 640 TCTGCTGGGTTGACCTCTGCACTTATCATTTGTTGTTTAACTTAAATGTTTCTGT 699

DB 601 TCTGCTGGGTTGACCTCTGCACTTATCATTTGTTGTTTAACTTAAATGTTTCTGT 660

QY 700 AAGAAAAGAAAGTATGAGTTGAGCTTTATTAACCTGGCCAACTTGCCTCCAGAGAGG 759

DB 661 AAGAAAAGAAAGTATGAGTTGAGCTTTATTAACCTGGCCAACTTGCCTCCAGAGAGG 720



QY 760 TTGGCAATGAGAGAGCAGTACGATTCCTCTGAGAGAAATATCTACACCATCGAGAG 819  
| | | | |  
Db 721 TTGGCAATGAGAGAGCAGTACGATTCCTCTGAGAGAAATATCTACACCATCGAGAG 780  
| | | | |  
QY 820 AACGATATGAGAGAGAGATTTCAATGAGTACTCTGCTTACGTCACAGCCGAGAGCA 879  
| | | | |  
Db 781 AACGATATGAGAGAGAGATTTCAATGAGTACTCTGCTTACGTCACAGCCGAGAGCA 840  
| | | | |  
QY 880 TCC 882  
| | | | |  
Db 841 TCC 843  
| | | | |  
RESULT 13  
BD157433 1772 bp DNA linear PAT 17-JAN-2003  
LOCUS BD157433  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD157433  
VERSION BD157433.1 GI:27863191  
KEYWORDS JP 2002191363-A/12276.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1772)  
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 12276 09-Jul-2002;  
HELIIX RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
PN JP 2002191363-A/12276  
PD 09-Jul-2002  
PF 28-Jul-2000 JP 2000280990  
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KYOKU  
PI SAITO,  
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
10,  
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS Location/Qualifiers  
1..1772  
Location/Qualifiers  
Source  
Location/Qualifiers  
1..1772  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 473 a 396 c 419 g 484 t  
ORIGIN  
Query Match 18.4%; Score 499.8; DB 6; Length 1772;  
Best Local Similarity 68.5%; Pred. No.1.7e-133;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTTACCCCTCAAGTGTCTCTGCTGC 76  
| | | | |  
Db 21 CTGACTTTTCTTTCGCAAGCTCAGTGTTCATGTTTCAATCTTCCCTTGAAGTGTCTGCTGC 80  
| | | | |  
QY 77 TGGCGCAACTACTCTTTCGCAAGCTCAGTGTTCAGGATGTTTAAAGTTTAAAGTTTAA 136  
| | | | |  
Db 81 TGGCGCTGCTACTACTTTCAGGATGTTTCAGGATGTTTAAAGTTTAAAGTTTAA 140  
| | | | |  
QY 137 ATGCTATCTGCTCCCTGCAAGTACTACTTACCTTACCTGAGCACTTGTGCTATGCT 196  
| | | | |  
Db 141 ATGCTATCTGCTCCCTGCTTCAACCCAGCGCGCCCGAGGAACCTGCTGCTGCT 200  
| | | | |  
QY 197 GGGGCAAGGATTTCTGCTTGTGTCACAGTACCAATGAGTTGCTGCAAACTGATGAA 256  
| | | | |  
Db 201 GGGGCAAGGAGCTGCTGCTGTTTGAATGTGCAACGTGTGCTCAGACTGATGAA 260  
| | | | |

QY 257 GAATGTGACATATCAGAAATCCAGAGATACAGCTTAAAGGGGATCTCAACAAAGAG 316  
| | | | |  
Db 261 GGAATGTGAATATTTCGACATC---CAGATTCGCTTAAATGGGATTTCCGCAAGAG 317  
| | | | |  
QY 317 ATGTGTCTCTGATCATTAAGAGATGTGACTGTGATGACCATGAGACCTTACTGTCAGGA 376  
| | | | |  
Db 318 ATGTGTCTCTGATCATTAAGAGATGTGACTGTGATGACCATGAGACCTTACTGTCGCGGA 377  
| | | | |  
QY 377 TACAGTTCCCTGCTTATGATTAATAAATTGAATGAATGAATTTAGACATCAAGAG 436  
| | | | |  
Db 378 TCCAAATCCAGGATATGATGATGAATAAATTGAATGAATTTAGACATTTGATCAACAG 437  
| | | | |  
QY 437 CCAAGTCACTCCAGCTCAGCTCAGCTCCAGGGGACTTACTACAGCTTCCCAAGAACCC 496  
| | | | |  
Db 438 CCAAGTCACTCCAGCTCAGCTCAGCTCCAGGGGACTTACTACAGCTTCCCAAGAGATGC 497  
| | | | |  
QY 497 TACCAAGAGAGAAATG---TTCAAGACACACAGACATGCTGACCTCCATATATACA 553  
| | | | |  
Db 498 TTACCAAGAGAGACATGCGCCACAGACAGACAGACATGCGGAGCTCCCTGATATATA 557  
| | | | |  
QY 554 ATGGAACAAATTTCCAC-----ATGGCTGATGAA 586  
| | | | |  
Db 558 ATCTAACAAATATCCACATTTGCGCAATGAGTTACGGGACTCTAGATTGGCCATGACT 617  
| | | | |  
QY 587 TTAGGACTCTGGAGAAACGATCAGACCTGCTATCCACATTTGAGTGGAGTCTGCTG 646  
| | | | |  
Db 618 TACGGAGCTCTGGAGCAACCATGAAATGAGCATCTACATGAGCAGGAGATCTGTG 677  
| | | | |  
QY 647 GGTGACCTCTGGACATTAATGCTTAAATCTTAAATGATTCCTGTAAGAAA 706  
| | | | |  
Db 678 GGTGACCTCTGGACATTAATGCTTAAATGATTCCTGTAAGAAA 737  
| | | | |  
QY 707 AGAAGTATGAGTTTAAAGCTTATTAACATGAGCCAACTTGTCTCCAGAGAGTTGGCAA 766  
| | | | |  
Db 738 AGAAGTATGAGTTTAAAGCTTATTAACATGAGCCAACTTGTCTCCAGAGATTTGGCAA 797  
| | | | |  
QY 767 ATGAGAGAGAGTACGATTCCTGAGAGAAATATCTACATGAGAGAGAACGAT 826  
| | | | |  
Db 798 ATGAGAGAGAGTACGATTCCTGAGAGAAATATCTACATGAGAGAGAACGAT 857  
| | | | |  
QY 827 ATGAGAGAGAGTACGATTCCTGAGAGAAATATCTACATGAGAGAGAACGAT 886  
| | | | |  
Db 858 ATGAGAGAGAGTACGATTCCTGAGAGAAATATCTACATGAGAGAGAACGAT 917  
| | | | |  
QY 887 CGCTCTGAGACTGCCACTTTAAAGC-----TGCCTTCAATTTGACTTTGG 935  
| | | | |  
Db 918 AACCTTTGGGTTGCTGCTTTCGATGATGATCAACCAACCTTATTTTGAAGCTTGT 977  
| | | | |  
QY 936 TATTTCCCTTTCGAGAAACATGATGATGATGATGATGATGATGATGATGATGATG 995  
| | | | |  
Db 978 GTTTGCTTTTTCGAGAAACATGATGATGATGATGATGATGATGATGATGATGATG 1035  
| | | | |  
QY 996 CCAAGCAGCAGTAAAGAGATTCAGATTTTCTGGAGATTAATTAATCAAGAGGAGATTC 1055  
| | | | |  
Db 1036 CCAAGCAGCAGTAAAGAGATTCAGATTTTCTGGAGATTAATTAATCAAGAGGAGATTC 1094  
| | | | |  
QY 1056 GACTGTATCTATCTACATTAAGATGATGATGATGATGATGATGATGATGATGATG 1112  
| | | | |  
Db 1113 GACTGTATCTATCTACATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 1169  
| | | | |  
QY 1155 AGTTACCAACCCAGAGACCTTATATGATGATGATGATGATGATGATGATGATGATG 1214  
| | | | |  
Db 1170 CATTAGGAATGTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1243  
| | | | |  
RESULT 14  
AK027334 1772 bp mRNA linear PRI 01-AUG-2002  
LOCUS AK027334



DEFINITION Homo sapiens CDNA FLJ14428 fis, clone HEMBA1006293.  
ACCESSION AK027334  
VERSION AK027334.1 GI:14041941  
KEYWORDS c11go capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,  
Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,  
Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,  
Makamatsu, A., Nakamura, Y., Nagahari, K., Masuno, Y. and Oshima, A.  
NEDO human cDNA sequencing project  
Unpublished  
TITLE 2 (bases 1 to 1772)  
JOURNAL Isogai, T. and Otsuki, T.  
REFERENCE Direct Submission  
AUTHORS Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; CDNA full insert sequencing;  
Research Association for Biotechnology; CDNA library construction;  
5'- & 3'-end one pass sequencing and clone selection; Helix  
Research Institute (supported by Japan Key Technology Center etc.)  
and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
FEATURES  
source Location/Qualifiers  
1. 1772  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HEMBA1006293"  
/cissue\_type="whole embryo, mainly head"  
/clone\_lib="HEMBA1"  
/dev\_stage="embryo, 10 weeks"  
/note="Cloning vector: pME185FL3"  
44..949  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55044.1"  
/db\_xref="GI:14041942"  
/translation="MESHLPDFCVLLLLLLLRSESEVRAEVCNATLPCTTPAA  
PENVLPVCKGKACVFEKGNVAVLRDEVDVNTSRVWLNDFRQDVLITENVL  
ADSGYICRDIQIPGIMNDEKFMKLVKIPAKVPAPTLQRDPFAAPRMLTTRGHGA  
EQTGLSLPDINLTQISTLANEIDRLANDLSDGATIRIGIYIAGICAGIALALI  
FCALLFKWYSHSKERIKONLSLISLANLPSGLANAVAEGRISBENITYIEENVVEVE  
PHEVYCVSRQPSQPLGCRFAMP"  
BASE COUNT 473 a 396 c 419 g 484 t  
ORIGIN  
Query Match 18.4%; Score 499.8; DB 9; Length 1772;  
Best Local Similarity 68.5%; Pred. No. 1.7e-133;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGGATTTCCCTCCCAAGTACTGATGTTTCAAGCTTACCTCAACTGTCTCTGCTG 76  
DB 21 CTGACTTTTCTTGTGCAAGCTCCATGTTTTCATCTTCCCTTGTGACTGTCTCTGCTG 80  
QY 77 TGCTGCACTACTACTTGCAAGTCACTGGAAGATGTTAAGTGAAGTGTGTAATA 136  
DB 81 TGTCTGTCTACTACTTCAAGGTCTCCAGAGTGAATACAGAGCGAGGTGGTCA 140  
QY 137 ATGCTATCTGCCCTGCACTTCACTTCACTTCACTTGGGACACTTGTGCTATGCT 196  
DB 141 ATGCTATCTGCCCTGCTTCTCAACCCGAGCGCCCGAGGAACCTGTGCTGCT 200  
QY 197 GGGGCAAGGATTTCTCTGCTGCTCAAGTACCAATGATGTTGCTCAAGTGAATA 256  
DB 201 GGGGCAAGGATTTCTCTGCTGCTGATGTTGCAACGCTGTGCTCAGAGTGAATA 260

QY 257 GAAATGTGCAATATCAGAAATCCAGCAATACCAGTAAAGGCCATCTCAACAAAGCG 316  
DB 261 GGAATGTGATTAATGACATC---CAGATATGCTGCTAAATGGGATTTCCCAAGGAG 317  
QY 317 ATGTGCTCTGATCAATAAGATGTGACTCTGATGATGACCATGGGACCTACTGTCGAGA 376  
DB 318 ATGTGCTCTGATCAATAAGATGTGACTCTGATGATGACCATGGGACCTACTGTCGAGA 377  
QY 377 TACAGTCCCTGCTGCTTATGAAATGATAAATAATGAACTGAATATGACATCAAGCAG 436  
DB 378 TCCAAATCCAGGCAATTAATGATGATAAATAATTAACCTGAAGTTGTCATCAACGAG 437  
QY 437 CCAAGGTCACTCCACTCAGATGCTCCCATGGGACCTTACTACAGCTTCCAAAGACC 496  
DB 438 CCAAGGTCACTCCACTCAGATGCTCCCATGGGACCTTACTACAGCTTCCAAAGATTG 497  
QY 497 TAAACAGGAGAGAAATGG---TTGAGAGACAGACACTGCTGACCTCCATTAATTA 553  
DB 498 TTACCAACAGGAGCAATGAGCCCAAGACAGACACAGACTGAGGAGCTCTCTGATATA 557  
QY 554 ATGAAACAAAATTTCCAC-----ATGGGCTGATGAA 586  
DB 558 ATCTAACAACAATATCCACATTTGGCCAAATGAGTTAGGGACTCTAGATTTGCCAATGACT 61  
QY 587 TTAAGGACTCTGAGAGAAAGATCAAGATCTATCACTTGGAGTGGAGTCTCTGCTG 646  
DB 618 TACGGACTCTGAGCAACATCAAGATGAGATCTCACTGAGACAGGATCTGTCTG 677  
QY 647 GGTGACCTGACCTATCATTTGCTTATCTTAAATGATGATTTCTGTAAGAAA 706  
DB 678 GGTGCTCTGCTCTTATCTTATCTTGGGCTTAAATTTCAATGATTTCTCAATAGCAAG 737  
QY 707 AAGATTTCAAGTTTGAAGCTTATTAATCACTGCACTGCTCCAGAGGCTTGCA 766  
DB 738 AAGAAATCAGAAATTAAGCTCATCTCTTGGCAACCTCCCTCCAGATTTGGCA 797  
QY 767 ATGCAAGGACATCAGAAATGCTGCTGAGAAATAATCTACACCATCGAGAGAACTGT 826  
DB 798 ATGCAAGGACATCAGAAATGCTGCTGAGAAATAATCTACACCATCGAGAGAACTGT 857  
QY 827 ATGAAATGAGAAATTCAAATGATGATCTGCTGACGTCAACAGCAGCAGCATCTGAC 886  
DB 858 ATGAAATGAGAGAGCCCAATGATGATTTATGCTATGTCAGAGCAGGAGCAACCTGAC 917  
QY 887 CCGCTCTGAGCTGCACTTTAAAGC-----TGGCTTCATTTCTGACTTTGG 935  
DB 918 AACCTTTGGGTTGTGCTTGGCAATGATGATCAACCACTTATTTTGAAGCTTGGT 977  
QY 936 TATTTCCCTTTTGAAATACTATGATATGCTGCTGCAACCTCATTTGAGGTTCTGA 995  
DB 978 GTTTGTCTTTTTCAGAACTATGAGCTGTGTCACTGCTGTT--TTGGAGGTTCTGT 1035  
QY 996 CCACAGCACTGAGAAAGAGTTTCCAGTTTCTGAGGATTAATTAATCAACAGGAGATT 1055  
DB 1036 CCACGCTATGAGAGAGATTTTCCATTTTCAAGAAATTA--TGACTCAATGAGGATTTG 1094  
QY 1056 GACTGTAATCTACTGATCAATGAATGCTCATTTTATCTCTGAGTTTCAGGG---ATCG 1112  
DB 1095 AACTGGAGCTGCACTGAATTAACAGGCAATGATGCTCTCTATTTTAAGCCAAAG 1154  
QY 1113 GATCTCCACCTCAGAGACT--TCATCATGAGCTTGAAGCTCACGCGGCTTT--CAT 1169  
DB 1155 AGTTACCAACCAAGAGACTGTTATTCATGATGATTTAGACTCAACGGGCTTTATATA 1214  
QY 1170 CATTAAGATGTTAGTGTGATGCTTTG 1198  
DB 1215 CACTAGGAATCTTGAACGAGGCTCTG 1243  
RESULT 15  
ARI43568 2236 bp DNA linear PAT 08-AUG-2001  
LOCUS ARI43568  
DEFINITION Sequence 23 from patent US 6204371.

ACCESSION ARI43568  
 VERSION ARI43568.1 GI:15104854  
 KEYWORDS  
 SOURCE  
 ORGANISM Unknown.  
 REFERENCES Unclassified.  
 1 (bases 1 to 2236)  
 AUTHORS Levinson, D. Adam.  
 TITLE Compositions and methods for the treatment and diagnosis of immune disorders  
 JOURNAL Patent: US 6204371-A 23 20-MAR-2001;  
 FEATURES Location/Qualifiers  
 source 1..2236  
 /organism="unknown"  
 BASE COUNT 586 a 518 c 525 g 606 t 1 others  
 ORIGIN  
 Query Match 18.4%; Score 499.8; DB 6; Length 2236;  
 Best Local Similarity 68.5%; Pred. No. 1.7e-133;  
 Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 91;  
 QY 17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCACTGTGCTGCTG 76  
 Db 19 CTGACTTTCTTCTGCAAGCTCCAGTGTTCACATCTCCCTTGACTGTGCTGCTG 78  
 QY 77 TGCGCACTACTACTGCAAGGTCATTTGAAGATGTTATAGGTTGAGGTTGATAAA 136  
 Db 79 TGTGCTCTCTACTCTTACAGAGTCTCAGAGTGAATACAGACCGAGGTGGTCA 138  
 QY 137 ATGCTATCTGCTCCGTCAGTGTACCTTACCTTACCTTGGGACACTGTGCTATG 196  
 Db 139 ATGCTATCTGCTCCGTCGTTTACACCCAGCGCCCGGAGAACTCGTGCCGTG 198  
 QY 197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGAGTTGCTCAAGTGA 256  
 Db 199 GGGGCAAGGAGCTGCTGCTGTGTTGATGTGCAAGTGTGCTCAGAGCTGAT 258  
 QY 257 GAAATGACATATCAGAAATCCAGAGATACCAAGTAAAGGCGATCTCAACAAAG 316  
 Db 259 GGGATGTAAATTTATGACATC--CAATATGCTAAATGAGGATTTCCGCAAG 315  
 QY 317 ATGTGCTCTGATCATTAAGATGATGACTGTGATGACCATGGACTTACTGTG 376  
 Db 316 ATGTGCTCTGATCATTAAGATGATGACTGTGATGACCATGGACTTACTGTG 375  
 QY 377 TACAGTTCCCTGCTTATGATGATTAATAAATTGAAGTGAATTAACATCAAG 436  
 Db 376 TCCAAATCCCAAGGCAATTAATGATGATGAATAATTAACCTGAAGTTGTC 435  
 QY 437 CCAAGGTACCTCCAGCTCAGAGTGGCCATGGGAGCTTACTACAGCTTCCAAAG 496  
 Db 436 CCAAGGTACCTCCAGCTCAGAGTGGCCATGGGAGCTTACTACAGCTTCCAAAG 495  
 QY 497 TAACCAAGGAGAGATGAG---TTGAGAGACAGACACTGGTGAACCTCCATAA 553  
 Db 496 TTACCAAGGAGAGATGAG---TTGAGAGAGACAGACACTGGGAGCTTCCCTG 555  
 QY 554 ATGAAACAAAATTTCCAC-----ATGGGCTGATGAAA 586  
 Db 556 ATCTAACACAAATATTCACATTTGCGCAATGAGTTACGGGACTCTAGATTG 615  
 QY 587 TTAAGGACTCTGAGAGAAACAGATCAGATCGTATCCAGATGGAGTGTGCTG 646  
 Db 616 TACGGGACTCTGAGAGAAACAGATCAGATCGTATCCAGATGGAGTGTGCTG 675  
 QY 647 GATTGACCTGGACCTTATCATTTGTTGTTTAACTTAATGATGATTTCTGTA 706  
 Db 676 GGGTGGCTCTGCTTATCTTGGGCTTTAAATTTTCAATGATTTCTATAGCAA 735  
 QY 707 AGAAGTTATGAGTTTGAAGCTTATTAACATGGCCAACTTGCTCCAGAGGTTG 766  
 Db 736 AGAAGTTATGAGTTTGAAGCTTATTAACATGGCCAACTTGCTCCAGAGTTG 795

QY 767 ATGCAAGAGAGAGTCAAGATTCGCTCTGAGAGAAATATCTACACCATGAGAGAG 826  
 Db 796 ATGCAAGTGAAGAGAGAGATTCGCTCAGAGAGAAATATCTACACCATGAGAGAG 855  
 QY 827 ATGAAATGAGAGATTCAAATGAGTACTACTGCTACGTCACAGGACGACGCTGAC 886  
 Db 856 ATGAAATGAGAGAGAGAGATTCGCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 915  
 QY 887 GCGCTGAGAGTGGCACTTTTAAAGG-----TGCCCTCAATTTCTGACTTGG 935  
 Db 916 AACCTTGGGTTGTGCTTGAATGCAATGATCAACCACTTATTTTGAAGCTTGT 975  
 QY 936 TATTTCCCTTTTGTGAGAAATATGATGATGATGATGATGATGATGATGATGAT 995  
 Db 976 GTTTTGTCTTTTGAAGAACTATGAGCTGTGTACCTGACCTGCTGTT--TTG 1033  
 QY 996 CCACAGCCACTGAGAGAGAGTTCAGTGTTCGAGAGATTAATTAATCAAGAGGAT 1055  
 Db 1034 CCACCTGCTATGAGAGAGAGTTCCTCATTTTCAAGAGATTA--TGACTCA 1092  
 QY 1056 GACTGTATCTGATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1112  
 Db 1093 AACTGGAGCTGCACTGAATTAACAGGCAATGATGATGATGATGATGATGATGAT 1152  
 QY 1113 GATCTCCACTGCAAGAGACT--TCAATCATGAGTGTGAGCTCACTGCTT--CAT 1169  
 Db 1153 AGTTACCAACCCAGAGAGCTTATCATGATGATGATGATGATGATGATGATGAT 1212  
 QY 1170 CATTGAGATGTTAGTGTGATGTTG 1198  
 Db 1213 CACTAGAAATCTTGAAGTGTGCTG 1241

Search completed: November 21, 2003, 23:34:26  
 Job time : 10139.7 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 11:32:52 ; Search time 747.069 Seconds  
(without alignments)  
9792.239 Million cell updates/sec

Title: US-10-004-633-8  
Perfect score: 2710  
Sequence: 1 ntgcgaccacccgcgcgcga.....aaaaaaaaagcgccgcgc 2710

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	21	AAAS1898
2	2706.6	99.9	2710	22	AAI70263
3	2706.6	99.9	2710	22	AAF82609
4	2706.6	99.9	2710	22	AAC90986
5	2706.6	99.9	2710	22	AAF23458
6	2706.6	99.9	2710	24	AB553316
7	2705	99.8	2710	25	ABQ77037
8	2684.6	99.1	2712	17	AAT38265

9	2624.6	96.8	2725	25	AB268328
10	844.4	31.2	862	25	AB268329
11	843	31.1	843	21	AAAS1918
12	843	31.1	843	22	AAI70255
13	843	31.1	843	22	AAF82615
14	843	31.1	843	22	AAC90992
15	843	31.1	843	22	AAF23464
16	843	31.1	843	25	ABQ77038
17	499.8	18.4	1772	22	AAH15441
18	499.8	18.4	2236	21	AAAS1904
19	499.8	18.4	2236	22	AAI70254
20	499.8	18.4	2236	22	AAF82616
21	499.8	18.4	2236	22	AAC90993
22	499.8	18.4	2236	22	AAF23465
23	499.8	18.4	2236	22	AB553322
24	499.8	18.4	2236	25	ABQ77040
25	499.8	18.4	1756	22	AA522714
26	499	18.4	1475	24	ABL90805
27	497.8	18.4	1800	22	AA522478
28	497.2	18.3	2237	17	AAI78267
29	485.8	17.9	1116	25	AB268338
30	485.8	17.9	1116	25	AB268339
31	456	16.8	903	21	AAAS1905
32	456	16.8	903	22	AAI70261
33	456	16.8	903	22	AAF82629
34	456	16.8	903	22	AAC92155
35	456	16.8	903	22	AAF23478
36	456	16.8	903	25	ABQ77041
37	384	14.2	1326	25	ABT33381
38	336.4	12.4	769	22	AAH04055
39	328.8	12.1	2091	20	AA35717
40	289	10.7	555	20	AA35716
41	263.4	9.7	529	20	AAV87937
42	231.6	8.5	1203	21	AA88799
43	219.8	8.1	438	20	AA41096
44	84.2	3.1	1079	25	AB268335
45	84.2	3.1	1080	25	AB268332

## ALIGNMENTS

RESULT 1	AAAS1898
ID	AAAS1898 standard; DNA; 2710 BP.
XX	AAAS1898;
AC	31-OCT-2000 (first entry)
DT	
XX	
DE	Murine T helper cell differentially expressed gene 200.
XX	
KW	T helper cell; differential expression; 200 gene; immunomodulator;
KW	anti-inflammatory; anti-arthritic; antibacterial; immunosuppressive;
KW	thymomimetic; anti-thyroid; anti-asthmatic; anti-allergic; antiviral;
KW	protozoacide; lymphocyte; modulator; gene therapy; ss.
OS	Mus sp.
XX	
FH	Key
FT	Location/Qualifiers
FT	40..885
FT	/*tag= a
FT	40..99
FT	/*tag= b
FT	100..882
FT	/*tag= c
PD	US6084083-A.
XX	
PD	04-JUL-2000.
XX	
PF	28-MAR-1997; 97US-0829525.
XX	

Nucleotide sequenc  
Nucleotide sequenc  
T helper cell diff  
Mouse 200 gene, pr  
Murine TH1 specifi  
Mouse nucleotide s  
DNA associated wit  
Murine TH1-associ  
Human cDNA sequenc  
Human T helper cel  
Human 200 gene, pr  
Human TH1 specific  
Human 200 gene nuc  
Human 200 gene. H  
DNA encoding human  
Human TH1-associat  
Human cDNA encodin  
Human polynucleoti  
Human cDNA encodin  
200 gene different  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Human T helper cel  
Human 200 gene, pr  
Human TH1 specific  
Human 200 gene nuc  
DNA associated wit  
Human TH1-associat  
NOVX DNA sequence  
Human cDNA clone (c  
cDNA encoding a pr  
EST clone F6708.  
Human SECK cDNA C1  
Human secreted pro  
Nucleotide sequenc  
Nucleotide sequenc

PR 01-MAR-1996: 96US-0609583.  
 PR 03-MAR-1995: 95US-0398533.  
 PR 07-JUN-1995: 95US-0487748.  
 XX  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI; 2000-464385/40.  
 DR P-PSDB; AAY97056.  
 XX  
 PT New isolated human 200 gene products or polypeptides, useful for  
 PT creating and diagnosing immune disorders, especially T helper  
 PT lymphocyte-related disorders  
 XX  
 PS Example; Fig 17A-D; 107pb; English.  
 XX  
 CC Genes which are differentially expressed within and among T helper (TH)  
 CC cells and TH cell subpopulations, e.g. TH0, TH1 and TH2 subpopulations,  
 CC can be used diagnostically or as targets for therapeutic intervention.  
 CC The polypeptides are useful for treating and diagnosing of immune  
 CC disorders, especially T lymphocyte-related disorders. These disorders  
 CC include chronic inflammatory diseases and disorders (e.g. Crohn's  
 CC disease, reactive arthritis, Lyme disease, Hashimoto's thyroiditis or  
 CC Grave's disease), or atopic conditions (e.g. asthma and allergy,  
 CC including allergic rhinitis or food allergies). Also included are  
 CC certain pathogen susceptibilities (e.g. leishmaniasis) and viral (e.g.  
 CC HIV) or bacterial (e.g. tuberculosis or lepromatous leprosy) infections.  
 XX  
 SQ Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;  
 Query Match 99.9%; Score 2706.6; DB 21; Length 2710;  
 Best Local Similarity 100.0%; Pred.No.0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 542 TCCATATAACATGGAACAAATTTCCACATGGCTGATGAATTAAGACCTCGAG 601  
 QY  
 DB 602 AACCATCAAGACTGCTATCCACATTGAGTGGAGTCTCTGCTGGGTGACCTGGCAC 661  
 QY  
 DB 602 AACCATCAAGACTGCTATCCACATTGAGTGGAGTCTCTGCTGGGTGACCTGGCAC 661  
 QY  
 DB 662 TTATCATTTGGTCTTAATCCCTTAATGGTATTCCTGTAAGAAAAGAAAGTATGAGT 721  
 QY  
 DB 662 TTATCATTTGGTCTTAATCCCTTAATGGTATTCCTGTAAGAAAAGAAAGTATGAGT 721  
 QY  
 DB 722 TGAGCCTTAATACAGTGGCCAACTTCCCTCAGAGAGGTGGCAATGACAGAGAGTCA 781  
 QY  
 DB 722 TGAGCCTTAATACAGTGGCCAACTTCCCTCAGAGAGGTGGCAATGACAGAGAGTCA 781  
 QY  
 DB 782 GGATTGCTCTGAGAGAAAATATCTACACCATGAGAGAACTATATGAGTGGAGATT 841  
 QY  
 DB 782 GGATTGCTCTGAGAGAAAATATCTACACCATGAGAGAACTATATGAGTGGAGATT 841  
 QY  
 DB 842 CAAATGAGTACTACGCTACGCTACGCAACAGCAGACAGCACTGACCGCTCGAGCTGCC 901  
 QY  
 DB 842 CAAATGAGTACTACGCTACGCTACGCAACAGCAGACAGCACTGACCGCTCGAGCTGCC 901  
 QY  
 DB 902 ACTTTTAAAGGCTGCTCATTTCTGATTTGATTTCCCTTTTGGAAAACTATGTC 961  
 QY  
 DB 902 ACTTTTAAAGGCTGCTCATTTCTGATTTGATTTCCCTTTTGGAAAACTATGTC 961  
 QY  
 DB 962 ATATGCTACTTGGCACTCATTGAGGTTCTGACACAGCCATGAGAAAAGATTCCA 1021  
 QY  
 DB 962 ATATGCTACTTGGCACTCATTGAGGTTCTGACACAGCCATGAGAAAAGATTCCA 1021  
 QY  
 DB 1022 GTTTCTGGGAAATTAATTAACACAGAGGAGATTCGATTAATCATGCTACATTGAAT 1081  
 QY  
 DB 1022 GTTTCTGGGAAATTAATTAACACAGAGGAGATTCGATTAATCATGCTACATTGAAT 1081  
 QY  
 DB 1082 GCTCATTTTATCCCTGATTTCAAGGATCGGATCTCCACATCAAGACCTTCAATCATG 1141  
 QY  
 DB 1082 GCTCATTTTATCCCTGATTTCAAGGATCGGATCTCCACATCAAGACCTTCAATCATG 1141  
 QY  
 DB 1142 CGTGTGAAGCTCCTGCTGCTTCAATTAAGAAATGTTATGTTGATGCTTTGAGA 1201  
 QY  
 DB 1142 CGTGTGAAGCTCCTGCTGCTTCAATTAAGAAATGTTATGTTGATGCTTTGAGA 1201  
 QY  
 DB 1202 CATPAGGTTTGGGATATCCGCAAGCTCCGCAACAGTATGAGGAGAAATTAAGGCTTA 1261  
 QY  
 DB 1202 CATPAGGTTTGGGATATCCGCAAGCTCCGCAACAGTATGAGGAGAAATTAAGGCTTA 1261  
 QY  
 DB 1262 GATAGGAAGTCCGCTTTTGTGATGTTGAAAAATCTTAAAGATTGATGCTTTCT 1321  
 QY  
 DB 1262 GATAGGAAGTCCGCTTTTGTGATGTTGAAAAATCTTAAAGATTGATGCTTTCT 1321  
 QY  
 DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1381  
 QY  
 DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1381  
 QY  
 DB 1382 AAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1441  
 QY  
 DB 1382 AAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1441  
 QY  
 DB 1442 AAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1501  
 QY  
 DB 1442 AAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGGCAATGCTTAACAGATAT 1501  
 QY  
 DB 1502 GGTCTGGAACAGTGAAGTGTATGTTGTAACAAGACAGAGAACTATATTTGATGTTGAC 1561  
 QY  
 DB 1502 GGTCTGGAACAGTGAAGTGTATGTTGTAACAAGACAGAGAACTATATTTGATGTTGAC 1561  
 QY  
 DB 1562 CCACATGAAAAAGCTAGGCTGTTAGAGATGTTTGAAGTCAAGAGATGAGAGATTAA 1621  
 QY  
 DB 1562 CCACATGAAAAAGCTAGGCTGTTAGAGATGTTTGAAGTCAAGAGATGAGAGATTAA 1621  
 QY  
 DB 1622 AGGCAACAAGATCCCGGGGGCTTGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1681  
 QY  
 DB 1622 AGGCAACAAGATCCCGGGGGCTTGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1681

QY 1682 GTCCACAAAGTCCCTGTCTCAGAGTATGAGTCTGAGTCTGAGGATGCGATGCGG 1741  
DB 1682 GTCCACAAAGTCCCTGTCTCAGAGTATGAGTCTGAGTCTGAGGATGCGGATGCGG 1741  
QY 1742 GGTTCCTCTCTCTCTCAGAGTATGAGTCTGAGTCTGAGGATGCGGATGCGG 1801  
DB 1742 GGTTCCTCTCTCTCTCAGAGTATGAGTCTGAGTCTGAGGATGCGGATGCGG 1801  
QY 1802 AC 1861  
DB 1802 AC 1861  
QY 1862 TCTCTATAACATGTATCTCTACAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
DB 1862 TCTCTATAACATGTATCTCTACAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
QY 1922 CATGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1981  
DB 1922 CATGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1981  
QY 1982 GAGTTCAGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2041  
DB 1982 GAGTTCAGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2041  
QY 2042 CGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2101  
DB 2042 CGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2101  
QY 2102 TGTCTAAAGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2161  
DB 2102 TGTCTAAAGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2161  
QY 2162 TGACAGAAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2221  
DB 2162 TGACAGAAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2221  
QY 2222 TAGTCAGCAAAATATCTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2281  
DB 2222 TAGTCAGCAAAATATCTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2281  
QY 2282 ACCTTAAGTTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2341  
DB 2282 ACCTTAAGTTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2341  
QY 2342 ATCTATAAGATCAAAATTAATAGACCCGAGGAAATTAATAGGAGAGAAAGT 2401  
DB 2342 ATCTATAAGATCAAAATTAATAGACCCGAGGAAATTAATAGGAGAGAAAGT 2401  
QY 2402 AGCTTTGTTCAAAAGATTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2461  
DB 2402 AGCTTTGTTCAAAAGATTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2461  
QY 2462 TAGTCAGCAAAATATCTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2521  
DB 2462 TAGTCAGCAAAATATCTTTGGGAGAGACGATGATGATGATGATGATGATGATGATGAT 2521  
QY 2522 TGCTTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2581  
DB 2522 TGCTTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2581  
QY 2582 TGAACAAATCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 2641  
DB 2582 TGAACAAATCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 2641  
QY 2642 TTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 2701  
DB 2642 TTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 2701  
QY 2702 GAGGAGGAGG 2710  
DB 2702 GAGGAGGAGG 2710

RESULT 2  
AA170263  
ID AA170263 standard; cDNA; 2710 BP.  
XX  
AC AA170263;  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Mouse 200 gene, preferentially expressed in TH1 cells.  
XX  
KW Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;  
KW receptor; differential expression; immune disorder; psoriasis;  
KW multiple sclerosis; insulin-dependent diabetes; antidiabetic;  
KW antipsoriatic; diagnosis; therapy; se.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 40..885  
FT /\*tag= a  
XX  
EN US6288218-B1.  
XX  
PD 11-SEP-2001.  
XX  
PF 25-SEP-1997; 97US-0937399.  
XX  
PR 01-MAR-1996; 96US-0609583.  
PR 03-MAR-1995; 95US-0398633.  
PR 07-JUN-1995; 95US-0487748.  
XX  
PA (LEVI/) LEVINSON D A.  
XX  
PI Levinson DA;  
XX  
DR WPI; 2001-647189/74.  
XX  
DR P-PSDB; AAM50225.  
XX  
PT Detecting 200 gene expression in a sample, useful for treating and  
PT diagnosing immune disorders, especially T lymphocyte-related disorders,  
PT comprises detecting the presence of a 200 gene product or an RNA  
PT encoding the 200 gene product  
XX  
PS Claim 2 (n); Fig 17A-D; 108pp; English.  
XX  
CC The present sequence is that of cDNA including the complete open  
CC reading frame of the mouse 200 gene. It encodes a 281-amino acid  
CC protein (see AAM50225). The cDNA was isolated using the 'T cell  
CC clone paradigm' and the 'transgenic T cell paradigm' approaches  
CC to the identification of genes that are differentially expressed  
CC in helper T cell (TH) subpopulations. The gene is expressed at  
CC a higher level in fully differentiated TH1 than in TH2 cells. It  
CC can be used diagnostically or as a target for therapeutic  
CC intervention for the treatment of immune disorders. A claimed  
CC method for diagnosing a TH cell subpopulation-related immune  
CC disorder involves detecting the level of a human 200 gene (see  
CC AAT70263) product, or an RNA encoding it, so that if the level  
CC differs from that in a control sample, the disorder is diagnosed.  
CC The disorder is especially a TH1 cell subpopulation-related immune  
CC disorder, such as multiple sclerosis, psoriasis or insulin-dependent  
CC diabetes (claimed). A method for detecting murine 200 gene  
CC expression is also claimed. In addition to the 200 gene, the  
CC invention provides other genes that are differentially expressed  
CC within and among TH cells and TH cell subpopulations and which can be  
CC used in methods for the diagnosis, prognosis, evaluation and treatment  
CC of TH cell subpopulation-related disorders, for the identification of  
CC subjects exhibiting a predisposition to such conditions, for  
CC monitoring patients undergoing clinical evaluation for the  
CC treatment of such disorders, and for monitoring the efficacy of  
CC compounds used in clinical trials. Other immune disorders that can  
CC be treated/diagnosed include Crohn's disease, reactive arthritis,  
CC Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis,





Db	1982	GAGTTCAGAGGACATGTGCCCCAACCTGCGACATCTTCTTTCACAAGAAAAGGTAC	204
OY	2042	CGCAAGCAGCTGCTGTCTGTAAAGAAAACCTGCGAAAAGGCAATTTGACTGTGTG	2101
Db	2042	CGCAAGCAGCCGTGCTGTGTAAAGAAAACCTGCGAAAAGGCAACTTGTGACTGTGTG	2101
OY	2102	TGCTCAAGGGGAACTGACATCGACGAACTTTCATTCCTGAGGGAACCTGAGCGTGTTC	2161
Db	2102	TGCTCAAGGGGAACTGACTCAGACACTTCTCATTTCTGAGGGAACCTGAGCGTGTTC	2161
OY	2162	TGACAGAAAGAACACCGGTGACTGAGACATACGAAGGAGAGCTTTGACGAATCTATA	2221
Db	2162	TGACAGAAAGAACACCGGTGACTGAGACATACGAAGGAGAGCTTTGACGAATCTATA	2221
OY	2222	TAGTCAGAGAAATATTTCTTTGGGAGAGACAGTGTGTAACCAATATGATTTCCAAAGCCGTGG	2281
Db	2222	TAGTCAGAGAAATATTTCTTTGGGAGAGAGTGTGTAACCAAAATGATTTTCCAAAGCCGTGG	2281
OY	2282	AACCTCAGTTTCATCTGGCTTACAGCTGCTGCCCCAGTGCCTTGATCTGTGCTGCTCC	2341
Db	2282	AACCTCAGTTTCATCTGGCTTACAGCTGCTGCCCCAGTGCCTTGATCTGTGCTGCTCC	2341
OY	2342	ATCTATAACAGAAATCAAAATTAATAATGACCCCGAGTGAATAATTAAGTACAGCAAAAGT	2401
Db	2342	ATCTATAACAGAAATCAAAATTAATAATGACCCCGAGTGAATAATTAAGTACAGCAAAAGT	2401
OY	2402	AGCTTTGTCAAAAGTTTTTTTGCATTGGGAGGAACTGTGTACATCAGAGACATCTGT	2461
Db	2402	AGCTTTGTCAAAAGTTTTTTTGCATTGGGAGGAACTGTGTACATCAGAGACATCTGT	2461
OY	2462	TAGTGAGAGACACAAAACCTGTGTACCGTTTTTTTCATGTATGAATTTTGTGTTAGGT	2521
Db	2462	TAGTGAGAGACACAAAACCTGTGTACCGTTTTTTTCATGTATGAATTTTGTGTTAGGT	2521
OY	2522	TGCTTCTTAGCTAGCTGTGAGGTCTTGCTTTCTTAAGTGGTATGAGAGGAGCACATC	2581
Db	2522	TGCTTCTTAGCTAGCTGTGAGGTCTTGCTTTCTTAAGTGGTATGAGAGGAGCACATC	2581
OY	2582	TAAACAAATCCATTAGAGATACACAGCTCATGTCAGAAAGGGAACCTAATCTCAATGCTT	2641
Db	2582	TAAACAAATCCATTAGAGATACACAGCTCATGTCAGAAAGGGAACCTAATCTCAATGCTT	2641
OY	2642	TTAAAGTATATAAACTGTACTGTGCAATCTTTGAGCATATAAAAAAAAAAAAAAAAAAG	2701
Db	2642	TTAAAGTATATAAACTGTACTGTGCAATCTTTGAGCATATAAAAAAAAAAAAAAAAAAG	2701
OY	2702	GGCGGCGCG 2710	
Db	2702	GGCGGCGCG 2710	

### RESULT 3

ID AAF82609 standard; cDNA; 2710 BP.

AC AAF82609;

DT 18-JUN-2001 (first entry)

Murine TH1 specific 200 gene CDNA.

Murine: T helper cell: TH1: TH2: immunomodulator:

KW antibacterial: T helper lymphocyte modulator: gene therapy

infection: ag.  
in specific gene; 200 gene; immune disorder; inflammation;  
KW

XX  
OS  
M18 an

XX	Key	Location/Qualifications
EH		

FT	40..884
ET	40..884
CDS	40..884

FT		/product= "200 gene product"
XX	US6204371-B1.	
XX	20-MAR-2001.	
PD		
XX	01-MAR-1996;	96US-0609583.
PE		
XX	03-MAR-1995;	95US-0396633.
PR	07-JUN-1995;	95US-0487748.
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Levinson DA;	
DR	WPI; 2001-272703/28.	
XX		
PT	New murine or human 200 genes and their corresponding polypeptides,	
PT	useful for treating or diagnosing immune disorders, especially T helper	
PT	lymphocyte-related disorders, e.g. inflammatory diseases or allergies	
PT	-	
PS	Claim 1; Fig 17; 109pp; English.	
XX		
PS	The present sequence is expressed at higher levels in T helper 1 (TH1)	
CC	cells than in T helper 2 (TH2) cells. The invention relates to an	
CC	isolated nucleic acid molecule, which comprises the full length murine	
CC	200 gene or full length human 200 gene nucleotide sequence. The nucleic	
CC	acids are useful for treating or diagnosing immune disorders,	
CC	especially T helper lymphocyte-related disorders, e.g. inflammatory	
CC	diseases (e.g. Crohn's disease), multiple sclerosis, Grave's disease,	
CC	contact dermatitis, psoriasis, asthma and allergies, or certain viral	
CC	(e.g. HIV) or bacterial (e.g. tuberculosis) infections.	
CC		
SQ	Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;	
	Query Match 99.9%; Score 2706.6; DB 22; Length 2710;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	2 GTGAGCCACGAGCGTCCGGATTTCCTCCCAAGTACTCATGTTTCAAGSTCTTAOCCTCA	61
Dd	2 GTCAACCACCGCTGCCGATTTCCCTCCCAAGTACTCATGTTCAGSTCTTACCTCA	61
OY	62 ACTGTGTCCTCTGCTGTGTGCACAATACTACTTGCAAGGTCATTTGAAGTGTAAAG	121
Dd	62 ACTGTGTCCTCTGCTGTGTGCACAATACTACTTGCAAGGTCATTTGAAGTGTAAAG	121
OY	122 TTGAGGTGTGTAATAATCCTTATCTGCTCGCATTAACACTTACCTTACCTTGAGCAC	181
Dd	122 TTGAGGTGTGTAATAATCCTTATCTGCTCGCATTAACACTTACCTTACCTTGAGCAC	181
OY	182 TTGGGCGCTATGTGCTGGGGCAAGGATCTGTCTTGTCACAGTATCCAATTGATGTC	241
Dd	182 TTGGGCGCTATGTGCTGGGGCAAGGATCTGTCTTGTCACAGTATCCAATTGATGTC	241
OY	242 TCAGAATCGATGAAGAATATGTGACATATCAGAATAATCCAGCAGATACCAAGTGAAGGCG	301
Dd	242 TCAGAATCGATGAAGAATATGTGACATATCAGAATAATCCAGCAGATACCAAGTGAAGGCG	301
OY	302 ATCTCAACAAAGAGATGTGTCTTGATCATTAAGAAATGTGACTGTGATGACCAATGGA	361
Dd	302 ATCTCAACAAAGAGATGTGTCTTGATCATTAAGAAATGTGACTGTGATGACCAATGGA	361
OY	362 CCTATCTGTGTGAGGATATACAGTTCCCTGTCTTATTAATGAATTAAGAACTGAAT	421
Dd	362 CCTATCTGTGTGAGGATATACAGTTCCCTGTCTTATTAATGAATTAAGAACTGAAT	421
OY	422 TAGACATCAAGCAGCCAAAGGTCACTCAGGTCAAGATCGCCATGAGGACTCTACTACAG	481
Dd	422 TAGACATCAAGCAGCCAAAGGTCACTCAGGTCAAGATCGCCATGAGGACTCTACTACAG	481
OY	482 CTTCCTCCAAAGACCTTAACACGAGAGAAATGTGTTAGAGACACAGACACTGTGTGACCC	541

Db 482 CTTTCCAGAAACCTTAACACCGAAGAAATGTTTCAGAGACACAGACTGAGTACC 541  
Qy 542 TCCATATATACATGAGAACAAAAATTTCCACATGGGCTGATGAAATTAAGACTCTGAG 601  
Db 542 TCCATATATACATGAGAACAAAAATTTCCACATGGGCTGATGAAATTAAGACTCTGAG 601  
Qy 602 AAACGATCAGAACTGCTATCCACATGGAGTGGAGCTCTGCTGGGTGACCTGGCAC 661  
Db 602 AAACGATCAGAACTGCTATCCACATGGAGTGGAGCTCTGCTGGGTGACCTGGCAC 661  
Qy 662 TTATCATGCTGTCTTATCTCTAAATGGATTCCTGTAGAGAAAAAGATTATGAGTT 721  
Db 662 TTATCATGCTGTCTTATCTCTAAATGGATTCCTGTAGAGAAAAAGATTATGAGTT 721  
Qy 722 TGAACCTTATTAACCTGCGCAACTTCACAGAGGTTGCAAAATGACAGAGACTCA 781  
Db 722 TGAACCTTATTAACCTGCGCAACTTCACAGAGGTTGCAAAATGACAGAGACTCA 781  
Qy 782 GGAATGCTCTGAGAAAAATATCTACACCATGAGAGAAAGTATATGAATGAGAGATT 841  
Db 782 GGAATGCTCTGAGAAAAATATCTACACCATGAGAGAAAGTATATGAATGAGAGATT 841  
Qy 842 CAAATGAGTACTGCTAGTACAGACAGAGAGGCTGACCGGCTCTGAGACTGCC 901  
Db 842 CAAATGAGTACTGCTAGTACAGACAGAGAGGCTGACCGGCTCTGAGACTGCC 901  
Qy 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGATATTTCTTTTGTGAAAACTATGTT 961  
Db 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGATATTTCTTTTGTGAAAACTATGTT 961  
Qy 962 ATATGTCATCTGGCACTTCATTTGAGGTTCTGACACAGGCTGAGAAAAAGTTCCA 1021  
Db 962 ATATGTCATCTGGCACTTCATTTGAGGTTCTGACACAGGCTGAGAAAAAGTTCCA 1021  
Qy 1022 GTTTTCGGGGATTAATTAATCAACAAGGGATTCGACTGATCACTGATCACTTGAAT 1081  
Db 1022 GTTTTCGGGGATTAATTAATCAACAAGGGATTCGACTGATCACTGATCACTTGAAT 1081  
Qy 1082 GCTCAATTTTATCCCTGAGTTTCAAGGATGATCTTCCACTCCAGAGCTTCAATCATG 1141  
Db 1082 GCTCAATTTTATCCCTGAGTTTCAAGGATGATCTTCCACTCCAGAGCTTCAATCATG 1141  
Qy 1142 CGTGTGAAGCTACCTGCTCTTCACTACTTGAAGATGTTAGTGTGATCTTTGAGA 1201  
Db 1142 CGTGTGAAGCTACCTGCTCTTCACTACTTGAAGATGTTAGTGTGATCTTTGAGA 1201  
Qy 1202 CATAGAGTTTGTGATATCCGCAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGGCTAA 1261  
Db 1202 CATAGAGTTTGTGATATCCGCAAAAGCTCTGAAACAGGTAGGGGGAATTAAGGGCTAA 1261  
Qy 1262 GATAGAGAGGTGCGGTCTTTGTTGATGTTGAAAAATCTTAAAGAGTTGTAGCTTTCT 1321  
Db 1262 GATAGAGAGGTGCGGTCTTTGTTGATGTTGAAAAATCTTAAAGAGTTGTAGCTTTCT 1321  
Qy 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGCAATGCTTAAACACATAT 1381  
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCCAGGTGCAATGCTTAAACACATAT 1381  
Qy 1382 AACTTGGGAACCTTAGGACAGAGGATGAATGTTCAAGGTACAGGCTATGCTGTA 1441  
Db 1382 AACTTGGGAACCTTAGGACAGAGGATGAATGTTCAAGGTACAGGCTATGCTGTA 1441  
Qy 1442 AGACTGTCTCAGCATCCAAAGACGAAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
Db 1442 AGACTGTCTCAGCATCCAAAGACGAAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
Qy 1502 GGTTCGGAACGTAGGATGCTATGTTGTAACAAGCAAGAGAACTTATATTTGATGATGACC 1561  
Db 1502 GGTTCGGAACGTAGGATGCTATGTTGTAACAAGCAAGAGAACTTATATTTGATGATGACC 1561  
Qy 1562 CCACATGAAAAAGCTAGGCTGTGAGACATGCTTGTAGACTCAAGATGAGAGGTA 1621  
Db 1562 CCACATGAAAAAGCTAGGCTGTGAGACATGCTTGTAGACTCAAGATGAGAGGTA 1621  
Qy 1622 AGGCAACAAGATCCCGGGGCTTCGTCGTCAGTCAAGCTTAAAGCTGATGTTCCA 1681  
Db 1622 AGGCAACAAGATCCCGGGGCTTCGTCGTCAGTCAAGCTTAAAGCTGATGTTCCA 1681  
Qy 1682 GTTCCAGAAAGTCCCTGTCTCAGAAAGACATGAGTGTATCTGCGGATGCTTCCATGGG 1741  
Db 1682 GTTCCAGAAAGTCCCTGTCTCAGAAAGACATGAGTGTATCTGCGGATGCTTCCATGGG 1741  
Qy 1742 GGTGTCTCTCTCTCTCAGAAAGACATGAGTGTATCTGCGGATGCTTCCATGGG 1801  
Db 1742 GGTGTCTCTCTCTCTCAGAAAGACATGAGTGTATCTGCGGATGCTTCCATGGG 1801  
Qy 1802 AC 1861  
Db 1802 AC 1861  
Qy 1862 TCTCTATATCATGTATCTCTTACAGACATCTCTCTGCTCTGTTAAGATAGTGGAG 1921  
Db 1862 TCTCTATATCATGTATCTCTTACAGACATCTCTCTGCTCTGTTAAGATAGTGGAG 1921  
Qy 1922 CATGGCAGAGCTCAGTAATTTATTTTCCAGACTCAGAAAGCTGAGCAGAACTGGA 1981  
Db 1922 CATGGCAGAGCTCAGTAATTTATTTTCCAGACTCAGAAAGCTGAGCAGAACTGGA 1981  
Qy 1982 GAGTTGAGAGCAGTGTGCTGCAAACTGCGCAGACTTCTTTCACAGAAAAAGTTACC 2041  
Db 1982 GAGTTGAGAGCAGTGTGCTGCAAACTGCGCAGACTTCTTTCACAGAAAAAGTTACC 2041  
Qy 2042 CCGAAGACCTGCTGTCTGTTAAAGAAACCTGGGAAAGCAACTTGACTGTTGG 2101  
Db 2042 CCGAAGACCTGCTGTCTGTTAAAGAAACCTGGGAAAGCAACTTGACTGTTGG 2101  
Qy 2102 TGTCTAAGGGGAATGACTCAGACAACTTCTCATCTCTGAGAAATCGAGCTGTTTC 2161  
Db 2102 TGTCTAAGGGGAATGACTCAGACAACTTCTCATCTCTGAGAAATCGAGCTGTTTC 2161  
Qy 2162 TGAACAGAGAACACCGGTGATCGGACATACAGAGGCAAGCTCTTGCAGCAATTTAT 2221  
Db 2162 TGAACAGAGAACACCGGTGATCGGACATACAGAGGCAAGCTCTTGCAGCAATTTAT 2221  
Qy 2222 TGTCTAAGGGGAATGACTCAGACAACTTCTCATCTCTGAGAAATCGAGCTGTTTC 2281  
Db 2222 TGTCTAAGGGGAATGACTCAGACAACTTCTCATCTCTGAGAAATCGAGCTGTTTC 2281  
Qy 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGCCTTGAATCTGTGCTGCC 2341  
Db 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGCCTTGAATCTGTGCTGCC 2341  
Qy 2342 ATCTATACAGAACTCAATTAATTAAGACCCGAGTGAATAATTAAGTGAAGAAAGT 2401  
Db 2342 ATCTATACAGAACTCAATTAATTAAGACCCGAGTGAATAATTAAGTGAAGAAAGT 2401  
Qy 2402 AGCTTGTCAAAAGATTTTGTGCAATGAGGAGCAACTGTATCAATCAAGGACATCTGT 2461  
Db 2402 AGCTTGTCAAAAGATTTTGTGCAATGAGGAGCAACTGTATCAATCAAGGACATCTGT 2461  
Qy 2462 TAGTAGAGACACCAAACTGTGTGATCCGTTTTTTCATGTATGAATTTGTGTTAGGT 2521  
Db 2462 TAGTAGAGACACCAAACTGTGTGATCCGTTTTTTCATGTATGAATTTGTGTTAGGT 2521  
Qy 2522 TGTCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGTGATAGAGAGGACATC 2581  
Db 2522 TGTCTTACCTAGCTGTGAGAGTCTGTGCTTTCTTAAGTGTGATAGAGAGGACATC 2581  
Qy 2582 TAAACAAATCCATTGAGATTAACAGCTCATGACAGAGGAGAACTAATCTCAATGTT 2641  
Db 2582 TAAACAAATCCATTGAGATTAACAGCTCATGACAGAGGAGAACTAATCTCAATGTT 2641  
Qy 2642 TTAAGTAAATTAACCTGATCTGCAAAAGTACTTGTAGCAATTAATTAATTAATTAAT 2701  
Db 2642 TTAAGTAAATTAACCTGATCTGCAAAAGTACTTGTAGCAATTAATTAATTAATTAAT 2701

Oy 2702 GCGGCGCC 2710  
Db 2702 GCGGCGCC 2710

RESULT 4  
AAC90986  
ID AAC90986 standard; DNA; 2710 BP.

AC AAC90986;

DT 20-MAR-2001 (first entry)

DE Mouse 200 gene nucleotide sequence SEQ ID NO:8.

XX Treatment; diagnosis; immune disorder; mast cell related disorder;  
XX T-helper lymphocyte-related disorder; ischaemic disorder;  
XX Identification; vasodilator; cardiast; antianginal; angina pectoris;  
XX ischaemic renal disease; myocardial ischaemia; myocardial infarction;  
XX cortical infarction; ischaemic injury; kidney transplant; de.

OS Mus musculus.

PN W0200073498-A1.

PD 07-DEC-2000.

PF 31-MAY-2000; 2000WO-US14986.

PR 02-JUN-1999; 99US-0324986.

PA (MILL-) MILLENNIUM PHARM INC.

PI Levinson DA, Lloyd CM, McCarthy SA;

DR WPI; 2001-016510/02.

XX P-PSDB; AAB50696.

PT Ameliorating a symptom of an ischemic disorder or injury in a mammal  
PT e.g. ischemic renal disease or myocardial ischemia. By administering a  
PT 200 gene product (SI), a nucleic acid encoding (SI) or an antibody  
PT directed against (SI) -

PS Claim 10; Fig 17; 309pp; English.

CC The present invention describes a method for ameliorating a symptom of  
CC an ischaemic disorder or injury in a mammal. The method comprises  
CC administering a 200 gene product, a nucleic acid encoding (SI) or an  
CC antibody directed against (SI). The method is useful for treating a  
CC symptom of an ischaemic disorder such as ischaemic renal disease or  
CC myocardial ischaemia (such as angina pectoris), myocardial or cortical  
CC infarction. The method is also useful for treating a symptom of an  
CC ischaemic injury occurring due to transplantation of a kidney. The  
CC present sequence is used in the exemplification of the present  
CC invention.

XX Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;

Query Match 99.9%; Score 2706.6; DB 22; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GTCCACCAAGCGGTCGATTTCCCTCCCAAGTACTGATTTTCAGGCTTACCTCA 61  
Db 2 GTCCACCAAGCGGTCGATTTCCCTCCCAAGTACTGATTTTCAGGCTTACCTCA 61  
Oy 62 ACTGATCTGCTGCTGCTGCAACTACTTGAAGGTCATTGGAAGATGTTAATAGG 121  
Db 62 ACTGATCTGCTGCTGCTGCAACTACTTGAAGGTCATTGGAAGATGTTAATAGG 121  
Oy 122 TTGAGGTTGTAATAATGCTATCTGCCCTGACGTTACCTACTACCTCTGGAGAC 181  
Db 122 TTGAGGTTGTAATAATGCTATCTGCCCTGACGTTACCTACTACCTCTGGAGAC 181

Db 122 TTGAGGTTGTAATAATGCTATCTGCCCTGACGTTACCTACTACCTCTGGAGAC 181  
Oy 182 TTGAGGTTGTAATAATGCTATCTGCCCTGACGTTACCTACTACCTCTGGAGAC 241  
Db 182 TTGAGGTTGTAATAATGCTATCTGCCCTGACGTTACCTACTACCTCTGGAGAC 241  
Oy 242 TCAGAACTGATGAAGAATGTCATATCAAAATCCAGAGATACCGAGTAAAGGCG 301  
Db 242 TCAGAACTGATGAAGAATGTCATATCAAAATCCAGAGATACCGAGTAAAGGCG 301  
Oy 302 ATCTCAAAAGAGATGTCCTGATCATAAAGATGTGACTCTGGATCAATGGGA 361  
Db 302 ATCTCAAAAGAGATGTCCTGATCATAAAGATGTGACTCTGGATCAATGGGA 361  
Oy 362 CCTACTGTCGAGATACAGTTCCTGCTTAAATGATTAATAAATAATTAAGCTGAAT 421  
Db 362 CCTACTGTCGAGATACAGTTCCTGCTTAAATGATTAATAAATAATTAAGCTGAAT 421  
Oy 422 TAGACATCAAGACAGCCAGGTCACCTCAGCTCAGACTGCCATGGGGACTCTACAG 481  
Db 422 TAGACATCAAGACAGCCAGGTCACCTCAGCTCAGACTGCCATGGGGACTCTACAG 481  
Oy 482 CTCTCCAAAGACCTTAACACGAGAGAAATGCTTCAAGACACAGCACTGTGACC 541  
Db 482 CTCTCCAAAGACCTTAACACGAGAGAAATGCTTCAAGACACAGCACTGTGACC 541  
Oy 542 TCCATTAATAAATGGAACAAATTTTCCATATGGGCTGATGAATTAAGATCTGGAG 601  
Db 542 TCCATTAATAAATGGAACAAATTTTCCATATGGGCTGATGAATTAAGATCTGGAG 601  
Oy 602 AAACGATCAGAACTGCTATCCACATGGAGGGAGTCTGCTGGGTTGACCCCTGGAC 661  
Db 602 AAACGATCAGAACTGCTATCCACATGGAGGGAGTCTGCTGGGTTGACCCCTGGAC 661  
Oy 662 TTATCATTTGTCCTTAATCTTAATGTAATTCCTGTAAGAAAAGATTAATGAGATT 721  
Db 662 TTATCATTTGTCCTTAATCTTAATGTAATTCCTGTAAGAAAAGATTAATGAGATT 721  
Oy 722 TGACCTTATTAACACTGGCCCACTTGCCTCAGAGAGGTTGGCAATTCAGAGAGTCA 781  
Db 722 TGACCTTATTAACACTGGCCCACTTGCCTCAGAGAGGTTGGCAATTCAGAGAGTCA 781  
Oy 782 GGATTCGCTGAGAGAAATATCTACACCATCGAGAGAAAGTATTAAGTGAATTT 841  
Db 782 GGATTCGCTGAGAGAAATATCTACACCATCGAGAGAAAGTATTAAGTGAATTT 841  
Oy 842 CAATGAGTACTACTGCTACGTCACAGCCAGCAGCCATCTGACCGCTGACTGCC 901  
Db 842 CAATGAGTACTACTGCTACGTCACAGCCAGCAGCCATCTGACCGCTGACTGCC 901  
Oy 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTCCCTTTTGGAAAATATG 961  
Db 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTCCCTTTTGGAAAATATG 961  
Oy 962 ATATGTCATCTGGACCTCATTTGAGAGGTTGACCAAGCAGCAGTGAAGAAAGTTCA 1021  
Db 962 ATATGTCATCTGGACCTCATTTGAGAGGTTGACCAAGCAGCAGTGAAGAAAGTTCA 1021  
Oy 1022 GTTTTCGGGATTAATTAATCAACAAGGATTTGACGTATATCATCTATCATTTGAAT 1081  
Db 1022 GTTTTCGGGATTAATTAATCAACAAGGATTTGACGTATATCATCTATCATTTGAAT 1081  
Oy 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGGAATCTCCACTCAAGAGACTTCAATCAT 1141  
Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGGAATCTCCACTCAAGAGACTTCAATCAT 1141  
Oy 1142 CGTGTGAAGCTCACTGCTGCTTATATCAATTAAGAAATGTTAGTGAATCTTTGGA 1201  
Db 1142 CGTGTGAAGCTCACTGCTGCTTATATCAATTAAGAAATGTTAGTGAATCTTTGGA 1201  
Oy 1202 CATAGAGGTTTGTGATATCCGAAAGCTCTGAAAGAGTGAAGGAGGATTAAGGCTTA 1261  
Db 1202 CATAGAGGTTTGTGATATCCGAAAGCTCTGAAAGAGTGAAGGAGGATTAAGGCTTA 1261

QY 1262 GATAGAGAGTGGGCTTTTGTGATGTGGAAAAATCTTAAGAAGTGGTACTTTTCT 1321  
 DB 1262 GATAGAGAGTGGGCTTTTGTGATGTGGAAAAATCTTAAGAAGTGGTACTTTTCT 1321  
 QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCGAGTGGCATATGCTTAACAGATAT 1381  
 DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAAGCGAGTGGCATATGCTTAACAGATAT 1381  
 QY 1382 AACTTGGGAACTTTAGGAGAGAGGATTAAGTTCAAGGTCAAGCCAGGCTATGCTGTA 1441  
 DB 1382 AACTTGGGAACTTTAGGAGAGAGGATTAAGTTCAAGGTCAAGCCAGGCTATGCTGTA 1441  
 QY 1442 AGACTCTTCAMATCCAAAGAGAAATTAACATGAGACAGCAGAGAGGCTGAGATGA 1501  
 DB 1442 AGACTCTTCAMATCCAAAGAGAAATTAACATGAGACAGCAGAGAGGCTGAGATGA 1501  
 QY 1502 GGCTGGAGACAGTGGGATTTGTATCAAGCAGAGGAATCTATATTTGATGCTGAGCC 1561  
 DB 1502 GGCTGGAGACAGTGGGATTTGTATCAAGCAGAGGAATCTATATTTGATGCTGAGCC 1561  
 QY 1562 CCAATGAAAAAGCTAGGCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGATTA 1621  
 DB 1562 CCAATGAAAAAGCTAGGCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGATTA 1621  
 QY 1622 AGGCAACAAGATCCCGGGGCTTGGGTGAGTCACTTACCTAGGTGCTGATTTCAA 1681  
 DB 1622 AGGCAACAAGATCCCGGGGCTTGGGTGAGTCACTTACCTAGGTGCTGATTTCAA 1681  
 QY 1682 GTCCACAAGATCCCGGTCTCAGAGTAAAGTGGCTGAGATCTGGGGCATGTCATGGG 1741  
 DB 1682 GTCCACAAGATCCCGGTCTCAGAGTAAAGTGGCTGAGATCTGGGGCATGTCATGGG 1741  
 QY 1742 GGTGTCTCTCTCTCAGAGAGACATGCAATGWCCTGACACACACACACACACAC 1801  
 DB 1742 GGTGTCTCTCTCTCAGAGAGACATGCAATGWCCTGACACACACACACACACAC 1801  
 QY 1802 ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATG 1861  
 DB 1802 ACACACACACACACACACACACACACACATGATGATGATGATGATGATGATG 1861  
 QY 1862 TCTCTTAACATATATCTTACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 DB 1862 TCTCTTAACATATATCTTACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 QY 1922 CATGCGAGAGCAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1981  
 DB 1922 CATGCGAGAGCAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1981  
 QY 1982 GAGTTCAAGAGCAGTGTGCCCCAACATGCGCAACTCTTTTCAACAAGAAAAAGTTAC 2041  
 DB 1982 GAGTTCAAGAGCAGTGTGCCCCAACATGCGCAACTCTTTTCAACAAGAAAAAGTTAC 2041  
 QY 2042 CGCAGAGAGCAGTGTGCTTAAGAGAAACCTGCGAAAGGAGAACTTTGACCTTTG 2101  
 DB 2042 CGCAGAGAGCAGTGTGCTTAAGAGAAACCTGCGAAAGGAGAACTTTGACCTTTG 2101  
 QY 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCAATTTCTGAGAGAAACTGAGCTTTTC 2161  
 DB 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCAATTTCTGAGAGAAACTGAGCTTTTC 2161  
 QY 2162 TGAACAAGAAACACCGGTACTGCGGACATACAGAGAGAGAGCTTTGCGCAATCTATA 2221  
 DB 2162 TGAACAAGAAACACCGGTACTGCGGACATACAGAGAGAGAGCTTTGCGCAATCTATA 2221  
 QY 2222 TAGTCAGCAAAATATTTCTTTGGAGGACAGTCTGACCAAAATGATTTTCAAGCCGGTGG 2281  
 DB 2222 TAGTCAGCAAAATATTTCTTTGGAGGACAGTCTGACCAAAATGATTTTCAAGCCGGTGG 2281  
 QY 2282 ACCTCAGTTTCATCTGCTTACAGCTGCTGCCAGTGCCTTGTATCTGTGCTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCATCTGCTTACAGCTGCTGCCAGTGCCTTGTATCTGTGCTGCTGCC 2341

QY 2342 ATCTATAACAGATCAATTAATTAAGACCCCGAGTGAAATTAATTAAGTGAAGAGGT 2401  
 DB 2342 ATCTATAACAGATCAATTAATTAAGACCCCGAGTGAAATTAATTAAGTGAAGAGGT 2401  
 QY 2402 AGCTTTGTTCAAGATTTTTTTCATGAGGAGACAACTGTGTATCATGAGAGACATCTGT 2461  
 DB 2402 AGCTTTGTTCAAGATTTTTTTCATGAGGAGACAACTGTGTATCATGAGAGACATCTGT 2461  
 QY 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTTTCATGATGAAATTTTGTGTTAGGT 2521  
 DB 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTTTCATGATGAAATTTTGTGTTAGGT 2521  
 QY 2522 TGCTTCTAGCTAGCTGTGAGAGTCCGGCTTCTTAAAGTGGTAAAGAGAGACATC 2581  
 DB 2522 TGCTTCTAGCTAGCTGTGAGAGTCCGGCTTCTTAAAGTGGTAAAGAGAGACATC 2581  
 QY 2582 TAAACAAATCCATTGAGATTAACAGCTCATGACAGAGAGAGAACTTAATCTCAATGTT 2641  
 DB 2582 TAAACAAATCCATTGAGATTAACAGCTCATGACAGAGAGAGAACTTAATCTCAATGTT 2641  
 QY 2642 TTAAGTAAATAAATCTGTACTGCGAAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
 DB 2642 TTAAGTAAATAAATCTGTACTGCGAAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
 QY 2702 GGGCGCCGC 2710  
 DB 2702 GGGCGCCGC 2710

RESULT 5  
 AAF23458  
 ID AAF23458 standard; DNA; 2710 BP.

AC AAF23458;  
 XX  
 DT 21-MAR-2001 (first entry)  
 XX  
 DE Murine 200 gene.  
 XX  
 KW Cysteine protease; immune disorder; T lymphocyte; Crohn's;  
 KW arthritis; diabetes; multiple sclerosis; viral infection; bacterial;  
 KW HIV; ds.  
 XX  
 OS Mus sp.  
 XX  
 PN US6156887-A.  
 XX  
 PD 05-DEC-2000.  
 XX  
 PF 03-OCT-1997; 97US-0939729.  
 XX  
 PR 01-MAR-1996; 96US-0609583.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX  
 DR WPI; 2001-101473/11.  
 XX  
 PT Novel polypeptide exhibiting cysteine protease activity, useful for  
 PT treating and diagnosing immune disorders, especially T  
 PT lymphocyte-related disorders, e.g. Crohn's disease, multiple sclerosis,  
 PT graft versus host disease or allergies -  
 XX  
 PS Examples; Fig 17; 107bp; English.  
 XX  
 CC The present invention relates to a novel cysteine protease. The  
 CC protein of the invention is useful for treating and diagnosing immune  
 CC disorders, especially T lymphocyte-related disorders. In particular,  
 CC the polypeptide is useful for treating or diagnosing T helper (TH) cell  
 CC or TH cell subpopulation-related disorders. These disorders include

CC Crohn's disease, reactive arthritis, Lyme disease, insulin-dependent  
CC diabetes, organ-specific autoimmunity, multiple sclerosis, Hashimoto's  
CC thyroiditis, Grave's disease, contact dermatitis, psoriasis, graft  
CC rejection, graft versus host disease, sarcoidosis, atopic (e.g. asthma  
CC or allergy), eosinophilia, conjunctivitis, glomerular nephritis, or  
CC helminthic (e.g. leishmaniasis), viral (e.g. HIV (human  
CC immunodeficiency virus)) or bacterial (e.g. tuberculosis or lepromatous  
CC leprosy) infections.

XX Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;

Query Match 99.9%; Score 2706.6; DB 22; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 2 GTGAGCCCAACGGGCTCCGGAATTCCTCCCAAGTACTCATGTTTTCAGGTCTTACCTCA 61
Db 2 GTGAGCCCAACGGGCTCCGGAATTCCTCCCAAGTACTCATGTTTTCAGGTCTTACCTCA 61
OY 62 ACTGTGTCTGTGCTGTGCAACTACTTGCAGAGTCAATGGAGTGGTTAAGG 121
Db 62 ACTGTGTCTGTGCTGTGCAACTACTTGCAGAGTCAATGGAGTGGTTAAGG 121
OY 122 TTGAGGTGGTAAATAATGCTTATCTGCTGCAATTACACTGTACTTACATCTGGACA 181
Db 122 TTGAGGTGGTAAATAATGCTTATCTGCTGCAATTACACTGTACTTACATCTGGACA 181
OY 182 TTGTGCCATATGTGCTGGGGCAAGGATTTGTCTCTTGGTCAAGTGTACATGAGTTC 241
Db 182 TTGTGCCATATGTGCTGGGGCAAGGATTTGTCTCTTGGTCAAGTGTACATGAGTTC 241
OY 242 TCAGAGATGATGAAGAATGTGACATATCAGAAATCCAGAGATTCAGCTAAAGGGG 301
Db 242 TCAGAGATGATGAAGAATGTGACATATCAGAAATCCAGAGATTCAGCTAAAGGGG 301
OY 302 ATCTCAACAAAGAGATGTGTCTGTGATCAATAAGATGTGACTGTGATGACATGGGA 361
Db 302 ATCTCAACAAAGAGATGTGTCTGTGATCAATAAGATGTGACTGTGATGACATGGGA 361
OY 362 CCTACTGTGCGAGTACGATTCCTGCTGTATGATATGATTAATAAATTGAATCTGAAT 421
Db 362 CCTACTGTGCGAGTACGATTCCTGCTGTATGATATGATTAATAAATTGAATCTGAAT 421
OY 422 TAGACATCAAAAGCAGCAAGTCACTCCAGCTCAGACTGCCATGGGGAGCTACTACAG 481
Db 422 TAGACATCAAAAGCAGCAAGTCACTCCAGCTCAGACTGCCATGGGGAGCTACTACAG 481
OY 482 CTTCCTCAAGAACCTTAACCAAGAGAGAAATGTTTCAGAGACACAGACACTGTGACCC 541
Db 482 CTTCCTCAAGAACCTTAACCAAGAGAGAAATGTTTCAGAGACACAGACACTGTGACCC 541
OY 542 TCCATTAATTAACATGGAACAAAATTTTCACATGGGCTGATGAAATTAAGGACTGTGAG 601
Db 542 TCCATTAATTAACATGGAACAAAATTTTCACATGGGCTGATGAAATTAAGGACTGTGAG 601
OY 602 AAACGATCAAGAACTGCTATCCATTTGAGAGTGGAGTCTCTGCTGGGTTGACCTTGGAC 661
Db 602 AAACGATCAAGAACTGCTATCCATTTGAGAGTGGAGTCTCTGCTGGGTTGACCTTGGAC 661
OY 662 TTATCATTTGTTCTTAATCCTTAATGTTATCTCTGTAAGAAAAGAAGTATGAGTT 721
Db 662 TTATCATTTGTTCTTAATCCTTAATGTTATCTCTGTAAGAAAAGAAGTATGAGTT 721
OY 722 TGAAGCTTATTAACATGCGCAACTGCTTCCAGAGGGTGGCAAAATGACAGAGAGTCA 781
Db 722 TGAAGCTTATTAACATGCGCAACTGCTTCCAGAGGGTGGCAAAATGACAGAGAGTCA 781
OY 782 GGAATTCGCTCTGAGAAAATATCTACACATGAGAGAAAGTATATGAGTGGAGATT 841
Db 782 GGAATTCGCTCTGAGAAAATATCTACACATGAGAGAAAGTATATGAGTGGAGATT 841
OY 842 CAATGAGTACTACTGTCTACGTCACAGGACAGGCAATCTGACCGGCTGTGACTGCC 901
Db 842 CAATGAGTACTACTGTCTACGTCACAGGACAGGCAATCTGACCGGCTGTGACTGCC 901
```

```
Db 842 CAATGAGTACTACTGTCTACGTCACAGGACAGGCAATCTGACCGGCTGTGACTGCC 901
OY 902 ACTTTAAAGGTGGCCTTCAATTTCTGACTTTGGATTTTCTTCTTGTGAAAACATATGCG 961
Db 902 ACTTTAAAGGTGGCCTTCAATTTCTGACTTTGGATTTTCTTCTTGTGAAAACATATGCG 961
OY 962 ATATGTCACTTGGCAACCTCAATTTGAGAGTTCGACACAGGCACTGAGAAAAGATTGCA 1021
Db 962 ATATGTCACTTGGCAACCTCAATTTGAGAGTTCGACACAGGCACTGAGAAAAGATTGCA 1021
OY 1022 GTTTTCTGGGATAATTAATCAACAGGGAGATTCAGCTGTAACTCATGCTACATGAAAT 1081
Db 1022 GTTTTCTGGGATAATTAATCAACAGGGAGATTCAGCTGTAACTCATGCTACATGAAAT 1081
OY 1082 GCTCATTTTATTCCTGAGTTTCAAGGATGGATTCCTCCACTCCAGAGACTTCAATCATG 1141
Db 1082 GCTCATTTTATTCCTGAGTTTCAAGGATGGATTCCTCCACTCCAGAGACTTCAATCATG 1141
OY 1142 CGTGTGAAGCTCACTGTGCTTTTATATCAATTAAGGATGTGATGTGATGCTTTGAGA 1201
Db 1142 CGTGTGAAGCTCACTGTGCTTTTATATCAATTAAGGATGTGATGTGATGCTTTGAGA 1201
OY 1202 CATAGAGTTTGTGTATATCCGCAAGCTCTCTGAACAGTATGGGGGAAATAAGGGCTTA 1261
Db 1202 CATAGAGTTTGTGTATATCCGCAAGCTCTCTGAACAGTATGGGGGAAATAAGGGCTTA 1261
OY 1262 CATAGAGTTTGTGTATATCCGCAAGCTCTCTGAACAGTATGGGGGAAATAAGGGCTTA 1261
Db 1262 CATAGAGTTTGTGTATATCCGCAAGCTCTCTGAACAGTATGGGGGAAATAAGGGCTTA 1261
OY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAGAGTGGCAATGCTTAACAGATAT 1381
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAGAGTGGCAATGCTTAACAGATAT 1381
OY 1382 AAATTGGGAACCTTAGGAGAGAGGTGATTAAGTTCAAGTCAAGGAGCTATGCTGTGA 1441
Db 1382 AAATTGGGAACCTTAGGAGAGAGGTGATTAAGTTCAAGTCAAGGAGCTATGCTGTGA 1441
OY 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATATGAGACAGGAGAGGCTGAGATGA 1501
Db 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATATGAGACAGGAGAGGCTGAGATGA 1501
OY 1502 GAGCTGGAAGTGAAGTGCATTTGTATCAAGACAGAGGAAATATATATTGATGTAGACC 1561
Db 1502 GAGCTGGAAGTGAAGTGCATTTGTATCAAGACAGAGGAAATATATATTGATGTAGACC 1561
OY 1562 CCAATTTGAAAAAGCTAGGCTGTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTTAA 1621
Db 1562 CCAATTTGAAAAAGCTAGGCTGTGTAGACATGCTTGTAGACTCAAGAGATGAGAGTTAA 1621
OY 1622 AGGCAACAAGATCCCGGGGCTTGGCTGACGTCAAGCTTAAGCTTAGGTGCTGATTCGA 1681
Db 1622 AGGCAACAAGATCCCGGGGCTTGGCTGACGTCAAGCTTAAGCTTAGGTGCTGATTCGA 1681
OY 1682 GTTCAACAGAGTCCCTGTCTCAAGTATGATGATGATGATGATGATGATGATGATGATG 1741
Db 1682 GTTCAACAGAGTCCCTGTCTCAAGTATGATGATGATGATGATGATGATGATGATGATG 1741
OY 1742 GGTGTCTCTCTCTCTCAAGAGAGATGACATATGATGATGATGATGATGATGATGATGAT 1801
Db 1742 GGTGTCTCTCTCTCTCTCAAGAGAGATGACATATGATGATGATGATGATGATGATGATGAT 1801
OY 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861
Db 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861
OY 1862 TCTCTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921
Db 1862 TCTCTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921
OY 1922 CATGGCAGAGAGTCAAGTAAATTTATTCACAGACTCAGAAAGCTGAGCAGAGGCTGGA 1981
Db 1922 CATGGCAGAGAGTCAAGTAAATTTATTCACAGACTCAGAAAGCTGAGCAGAGGCTGGA 1981
```

QY 1982 GAGTTGAGGAGCACTGTGCCCAACTGTCACATCTTCTTACACAGAAAAAGTTACC 2041  
 DB 1982 GAGTTGAGGAGCACTGTGCCCAACTGTCACATCTTCTTACACAGAAAAAGTTACC 2041  
 QY 2042 CGAAGCAGCCTGCTGTGTAAAGAAACCTGTGAAAGCAACTTGTGCTGTGTG 2101  
 DB 2042 CGAAGCAGCCTGCTGTGTAAAGAAACCTGTGAAAGCAACTTGTGCTGTGTG 2101  
 QY 2102 TGTCAAGGGGAACTGACTCAGCAACTTCTCCATTCCTGAGAAACTGTAGCTGTTC 2161  
 DB 2102 TGTCAAGGGGAACTGACTCAGCAACTTCTCCATTCCTGAGAAACTGTAGCTGTTC 2161  
 QY 2162 TGAAGAAAGCAACCGGTGACTGAGCAATACAGAGGAGAGCTTGTGAGCAATCTATA 2221  
 DB 2162 TGAAGAAAGCAACCGGTGACTGAGCAATACAGAGGAGAGCTTGTGAGCAATCTATA 2221  
 QY 2222 TGTGAGCAAAATATCTTGTGGAGGAGCAAGTGTGACCAAAATGATTTCCAGCCGCTG 2281  
 DB 2222 TGTGAGCAAAATATCTTGTGGAGGAGCAAGTGTGACCAAAATGATTTCCAGCCGCTG 2281  
 QY 2282 ACCTCAGTTTCATCTGCTTACAGCTGCTGCCAGTGTGCTGTGCTGTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCATCTGCTTACAGCTGCTGCCAGTGTGCTGTGCTGTGCTGCC 2341  
 QY 2342 ATCTAATACAGATCAATTAATATAGACCCCGAGTGAATAATTAAGTACAGCAAGGT 2401  
 DB 2342 ATCTAATACAGATCAATTAATATAGACCCCGAGTGAATAATTAAGTACAGCAAGGT 2401  
 QY 2402 AGCTTGTTCAAAGATTTTTTTCATTTGGGAGCAACTGTGTATCATACAGAGCATCTGT 2461  
 DB 2402 AGCTTGTTCAAAGATTTTTTTCATTTGGGAGCAACTGTGTATCATACAGAGCATCTGT 2461  
 QY 2462 TAGTGAGACACCAAAACCTGTGTGACCGTTTTCATGTATGATTTGTGTAGGT 2521  
 DB 2462 TAGTGAGACACCAAAACCTGTGTGACCGTTTTCATGTATGATTTGTGTAGGT 2521  
 QY 2522 TGTCTTACGTAAGCTGTGAGGTCTGTGCTTCTTGTGAGTGTGAGAGGAGACATC 2581  
 DB 2522 TGTCTTACGTAAGCTGTGAGGTCTGTGCTTCTTGTGAGTGTGAGAGGAGACATC 2581  
 QY 2582 TAAACAAATCCATTAAGATTAACAGCTCTCATGAGAGGAAACTTAATCTAAATGTT 2641  
 DB 2582 TAAACAAATCCATTAAGATTAACAGCTCTCATGAGAGGAAACTTAATCTAAATGTT 2641  
 QY 2642 TTAAGTAATTAAGCTGTGAGCAAGTACTTGTGAGCATTAATAAAAAAAAAAAG 2701  
 DB 2642 TTAAGTAATTAAGCTGTGAGCAAGTACTTGTGAGCATTAATAAAAAAAAAAAG 2701  
 QY 2702 GGGGGGGC 2710  
 DB 2702 GGGGGGGC 2710  
 RESULT 6  
 ID ABS53316 standard, cDNA, 2710 BP.  
 AC ABS53316;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Mouse Th1 specific 200 gene.  
 XX  
 TH1; TH2; se; antiinflammatory; antidiabetic; antithyroid; atopy;  
 KM antiaesthetic; antiallergic; virucide; antiarthritic; dermatological;  
 KM antipsoiatic; nephrotoxic; immunosuppressive; immune response;  
 KM asthma; allergy; allergic rhinitis; viral infection; thyroiditis;  
 KM inflammatory disease; Crohn's disease; arthritis; diabetes; mouse;  
 KM dermatitis; psoriasis; glomerular nephritis; autoimmunity; gene;  
 KM graft rejection.  
 XX  
 OS Mus sp.

XX Key Location/Qualifiers  
 FH CDS 40..885  
 FT /tag= a  
 FT /product= "200 gene protein"  
 PN US6414117-B1.  
 PD 02-JUL-2002.  
 PF 12-MAY-1999; 99US-0310367.  
 PR 01-MAR-1996; 96US-0609583.  
 PR 28-MAR-1997; 97US-0829525.  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Levinson DA;  
 DR WPI; 2002-641576/69.  
 DR P-PSDB; ABG32767.  
 XX  
 PT Novel polypeptides which are differentially expressed within and among  
 PT T helper cells and cell populations, useful for treating immune  
 PT disorders, especially T helper cell subpopulation-related disorders -  
 XX Example; Fig 17; 109pp; English.  
 PS  
 CC This invention relates to a novel isolated polypeptide which is a T.  
 CC helper (Th1) cell subpopulation specific gene product. The invention  
 CC also discloses other Th1 or Th2 specific gene products. The proteins of  
 CC the invention may have antiinflammatory, antidiabetic, antithyroid,  
 CC antiaesthetic, antiallergic, virucide, antiarthritic, dermatological,  
 CC antipsoiatic, nephrotoxic and immunosuppressive activities and may be  
 CC used as a regulator of the immune response. The proteins of the  
 CC invention are also useful to reduce the level of Th2 cell activity for  
 CC treating Th1 cell subpopulation-related disorders including atopic  
 CC conditions, such as asthma and allergy including allergic rhinitis, the  
 CC effects of pathogen, including viral infection, chronic inflammatory  
 CC diseases such as Crohn's disease, arthritis, diabetes, thyroiditis,  
 CC dermatitis, psoriasis, glomerular nephritis, organ-specific  
 CC autoimmunity, graft rejection and graft versus host disease. The  
 CC present sequence represents the DNA sequence of gene 200 isolated from  
 CC a Th1/Th2 differential display library, this DNA is specifically  
 CC expressed in Th1 cells.  
 CC  
 SQ Sequence 2710 BP; 789 A; 592 C; 644 G; 678 T; 7 other;  
 Query Match 99.9%; Score 2706.6; DB 24; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GTGACCCAGGGGGTCCGATTCCTCCCAAGTACTGATTTTCAGGCTTACCTCA 61  
 DB 2 GTGACCCAGGGGGTCCGATTCCTCCCAAGTACTGATTTTCAGGCTTACCTCA 61  
 QY 62 ACTGTCTCTGCTGTGCTGCAACTACTTGCAGAGTCTATGGAAGTGTATPAG 121  
 DB 62 ACTGTCTCTGCTGTGCTGCAACTACTTGCAGAGTCTATGGAAGTGTATPAG 121  
 QY 122 TTGAGGTGTAAATGCTATCTGCTGCAAGTTACCTTACTAATCTGGAGAC 181  
 DB 122 TTGAGGTGTAAATGCTATCTGCTGCAAGTTACCTTACTAATCTGGAGAC 181  
 QY 182 TTGTGCTATGTGCTGGGCAAGGATTTGCTTGTGCAAGTATCAATGAGTTGC 241  
 DB 182 TTGTGCTATGTGCTGGGCAAGGATTTGCTTGTGCAAGTATCAATGAGTTGC 241  
 QY 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGAGATCCAGTAAAGGGG 301  
 DB 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGAGATCCAGTAAAGGGG 301



OY	302	ATCTCAACAAAGAGATGTGTCCTGATCATTAAGAATGTGACTGTGATGACCATGGGA	361
Db	302	ATCTCAACAAAGAGATGTGTCCTGATCATTAAGAATGTGACTGTGATGACCATGGGA	361
OY	362	CTTACTGTGCGAGGATACAGTTCCCTGGCTTAATGAATGATTAATAAATTGAACATGAAT	421
Db	362	CTTACTGTGCGAGGATACAGTTCCCTGGCTTAATGAATGATTAATAAATTGAACATGAAT	421
OY	422	TAGACATCAAAAGCAGCCAAAGTCACTCCAGCTCAGACTGCCCCATGGGACTCTACACG	481
Db	422	TAGACATCAAAAGCAGCCAAAGTCACTCCAGCTCAGACTGCCCCATGGGACTCTACACG	481
OY	482	CTTCTCCCAAGAACCTTAACCAAGAGAGAAATGTTTCAAGACACAGACACTGTGTACCC	541
Db	482	CTTCTCCCAAGAACCTTAACCAAGAGAGAAATGTTTCAAGAGACACAGACACTGTGTACCC	541
OY	542	TCCATTAATAACAATGAACAAATAATTCACAATGGGCTGATGAATTAAGACTCTGGAG	601
Db	542	TCCATTAATAACAATGAACAAATAATTCACAATGGGCTGATGAATTAAGACTCTGGAG	601
OY	602	AAACGATCAGAACTGCTATCCACATTTGAGTGGAGTCTGTGCTGGTTGACCTCGGAC	661
Db	602	AAACGATCAGAACTGCTATCCACATTTGAGTGGAGTCTGTGCTGGTTGACCTCGGAC	661
OY	662	TTATCATTTGGTGTCTTAAATCCTTAAATGGTATTCCTGTAAAGAAAAAGAAATTAAGATT	721
Db	662	TTATCATTTGGTGTCTTAAATCCTTAAATGGTATTCCTGTAAAGAAAAAGAAATTAAGATT	721
OY	722	TGAGCCTTATTACACTGGGCAACTGGCCTCCAGAGAGGGTGGCAANTGACGAGGACATCA	781
Db	722	TGAGCCTTATTACACTGGGCAACTGGCCTCCAGAGAGGGTGGCAANTGACGAGGACATCA	781
OY	782	GGATTGCGCTGTGAGGAAAAATATCTACACCATCGAGAGAGACGTATATGAATGAGAAATT	841
Db	782	GGATTGCGCTGTGAGGAAAAATATCTACACCATCGAGAGAGACGTATATGAATGAGAAATT	841
OY	842	CAANTGAATACATACGTGCTACGTCAACAGCAGACAGCCATCTGTGACCGCTCTGTGACTGCC	901
Db	842	CAANTGAATACATACGTGCTACGTCAACAGCAGACAGCCATCTGTGACCGCTCTGTGACTGCC	901
OY	902	ACTTTTAAAGGCTCGCTTCATTTCTGACTTGTGTATTTCCCTTTKTGGAAAAATATGNG	961
Db	902	ACTTTTAAAGGCTCGCTTCATTTCTGACTTGTGTATTTCCCTTTKTGGAAAAATATGNG	961
OY	962	ATATGTCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAAGATTCCA	1021
Db	962	ATATGTCACTTGGCAACCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAAGATTCCA	1021
OY	1022	GTTTTCTTGGGATTAATTAATCTCACAAAGGGGATTTGCACTGTAACTCATGCTACATGAAT	1081
Db	1022	GTTTTCTTGGGATTAATTAATCTCACAAAGGGGATTTGCACTGTAACTCATGCTACATGAAT	1081
OY	1082	GCTCCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCACTCCAGAGACTTCATCATG	1141
Db	1082	GCTCCATTTTATCCCTGAGTTTCAAGGATCGGATCTCCCACTCCAGAGACTTCATCATG	1141
OY	1142	CGTGTGGAAGCTCACTCGTCTTTATACATTAGGAATGTTAGTGTGATGCTTTTGAGA	1201
Db	1142	CGTGTGGAAGCTCACTCGTCTTTATACATTAGGAATGTTAGTGTGATGCTTTTGAGA	1201
OY	1202	CATAGAGTTTGTGTATATCCGCAAGAGCTCTGAAACAGTATGGGGGAAATTAAGGGCTAA	1261
Db	1202	CATAGAGTTTGTGTATATCCGCAAGAGCTCTGAAACAGTATGGGGGAAATTAAGGGCTAA	1261
OY	1262	GATAGGAAGTGTGGTCTTTGTTGATGTTGAAAACTTTAAGAAATGTTGTAGCTTTTCT	1321
Db	1262	GATAGGAAGTGTGGTCTTTGTTGATGTTGAAAACTTTAAGAAATGTTGTAGCTTTTCT	1321
OY	1322	AGAGATTTCTGACTTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT	1381
Db	1322	AGAGATTTCTGACTTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT	1381

QY	1382	AACTTGGGAACCTTAGCGAGGAGGGGTATAGTTCAAGGTACAGCGGCGTATCTGGTA	14411
Db	1392	AACCTGGGAACCTTAGCGAGGAGGGGTATAGTTCAAGGTACAGCGGCGTATCTGGTA	14411
QY	1442	AGACTGTCTCMCAATCCAAAGACGAAAAATAACAATAGAGCAGCAGGAGGCTTGAGATGA	15011
Db	1442	AGACTGTCTCMCAATCCAAAGACGAAAAATAACAATAGAGCAGCAGGAGGCTTGAGATGA	15011
QY	1502	GAGCTCGACAGTGAAGGTGCATTGTGTACAGACACAGGAGATCTATATTGATCGTAGACC	15611
Db	1502	GAGCTCGACAGTGAAGGTGCATTGTGTACAGACACAGGAGATCTATATTGATCGTAGACC	15611
QY	1552	CCACATGAAAAAGCTAAGGCTCTGTATAGACATCTGTTTATGACTCAAGAGATGAGAGGTAA	16211
Db	1552	CCACATGAAAAAGCTAAGGCTCTGTATAGACATCTGTTTATGACTCAAGAGATGAGAGGTAA	16211
QY	1622	AGGCACACACAGATCCCCGGGGGCTTGCGTGACAGTACCTTAAGGTCGTGAGTCCAA	16811
Db	1622	AGGCACACACAGATCCCCGGGGGCTTGCGTGACAGTACCTTAAGGTCGTGAGTCCAA	16811
QY	1682	GTCCACAAAGAGTCCCTGTCTCMAGTAAGATGAGCTGATCTGCGGCATGTCATGGG	17411
Db	1682	GTCCACAAAGAGTCCCTGTCTCMAGTAAGATGAGCTGATCTGCGGCATGTCATGGG	17411
QY	1742	GGTGTCTCTCTCTCTCAGAAAGACATGCACATGACCCTGCACACACACACACACAC	18011
Db	1742	GGTGTCTCTCTCTCTCAGAAAGACATGCACATGACCCTGCACACACACACACACAC	18011
QY	1802	AACACACACACACACACACACACACACACATGAAATGAAAGTTCTCTGTGCTCTGTAAC	18611
Db	1802	AACACACACACACACACACACACACACACATGAAATGAAAGTTCTCTGTGCTCTGTAAC	18611
QY	1862	TCTCTATTAACATGTATCTCTACAGGACTCTCTCTGCTCTGTTTAAGACATGAGTGGAG	19211
Db	1862	TCTCTATTAACATGTATCTCTACAGGACTCTCTCTGCTCTGTTTAAGACATGAGTGGAG	19211
QY	1922	CATGCGACAGCAGTCCAGTAATTTATTCCAGCACTCAGAAAGGCTGAGCAGAGCGTGA	19811
Db	1922	CATGCGACAGCAGTCCAGTAATTTATTCCAGCACTCAGAAAGGCTGAGCAGAGCGTGA	19811
QY	1982	GAGTTACAGGACACATGAGCCCAACATGCGCAGACTCTTCTTACACAAAGAAAGATTACC	20411
Db	1982	GAGTTACAGGACACATGAGCCCAACATGCGCAGACTCTTCTTACACAAAGAAAGATTACC	20411
QY	2042	CGCAAGCAGCCTGCTGTCTGTAAAAAGGAAACCTGCGAAAGGCAAACTTTGACTGTTGTG	21011
Db	2042	CGCAAGCAGCCTGCTGTCTGTAAAAAGGAAACCTGCGAAAGGCAAACTTTGACTGTTGTG	21011
QY	2102	TGCTCAAGGGGAACTGACTCAGACAACTTCTCAATCTCTGAGGAAACTGAGGCTGTTTC	21611
Db	2102	TGCTCAAGGGGAACTGACTCAGACAACTTCTCAATCTCTGAGGAAACTGAGGCTGTTTC	21611
QY	2162	TGACACAGAAACCAACCGGTGACTGCGGACATCGAAGGACAGGCTCTTGACAAATCTATA	22211
Db	2162	TGACACAGAAACCAACCGGTGACTGCGGACATCGAAGGACAGGCTCTTGACAAATCTATA	22211
QY	2222	TAGTACGAAATAATTTCTTTGGGAGGACAGTCGTCAACCAATTTGATTTCCAAAGCCGGTG	22811
Db	2222	TAGTACGAAATAATTTCTTTGGGAGGACAGTCGTCAACCAATTTGATTTCCAAAGCCGGTG	22811
QY	2282	ACCTCAGTTTCACTGCGCTTACAGCTGCTGCCAGTGCCCTTATATCTGTGCTGGCTGCC	23411
Db	2282	ACCTCAGTTTCACTGCGCTTACAGCTGCTGCCAGTGCCCTTATATCTGTGCTGGCTGCC	23411
QY	2342	ATCTATAACAGATCAAAATTAAATAGACCCCGAGTGAATAATTTAAGTGAACAAAGGT	24011
Db	2342	ATCTATAACAGATCAAAATTAAATAGACCCCGAGTGAATAATTTAAGTGAACAAAGGT	24011
QY	2402	AGCTTTGTTCAAAAGATTTTTTTTGAATGGGAGGAACTGTGTACATCAGAGAGATCTGT	24611
Db	2402	AGCTTTGTTCAAAAGATTTTTTTTGAATGGGAGGAACTGTGTACATCAGAGAGATCTGT	24611
QY	2462	TAGTGAGACACAAACCTGTGATCCGTTTTTTCATGTATGATATTTGTTGTATAGGT	25211



DB 2462 TAGTAGAGACACAAAACCTGTGTACCGCTTTTTCATGATTAATTTTGTGTAGT 2521  
QY 2522 TGTCTTACCTAGCTGTGTGAGGCTCTGCTTTCTTAAAGTGGTATGGAAGGAGACCATC 2581  
DB 2522 TGTCTTACCTAGCTGTGTGAGGCTCTGCTTTCTTAAAGTGGTATGGAAGGAGACCATC 2581  
QY 2582 TAAACAAATCCATTAAGATTAACAGCTCTATGACAGAGGAAACTAATCTCAATGTT 2641  
DB 2582 TAAACAAATCCATTAAGATTAACAGCTCTATGACAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTAAGCTAATTAATAGTACTGCTGCAAGTACTTGTGACATTAATAAAAAAAAAAAG 2701  
DB 2642 TTAAGCTAATTAATAGTACTGCTGCAAGTACTTGTGACATTAATAAAAAAAAAAAG 2701  
QY 2702 GCGGCGCGC 2710  
DB 2702 GCGGCGCGC 2710  
RESULT 7  
ID ABQ77037 standard; cDNA; 2710 BP.  
AC ABQ77037;  
DT 09-APR-2003 (first entry)  
XX Murine TH1-associated 200 gene.  
DE  
XX T-cell receptor; TH; T helper cell; gene; 103 gene; TH2 cell; murine;  
KM TH2 cell marker; TH1 cell; protozoicide; antibacterial; vitruicide;  
XX immunosuppressive; antiinflammatory; antirheumatic; antidiabetic;  
KM neuroprotective; dermatological; antichryoid; antiporiatic; helminthic;  
XX nephrotoxic; antiaesthetic antiallergic; CD8 agonist; CD4 agonist;  
KM interleukin agonist; bacterial; viral infection; immune disorder;  
XX Crohn's disease; reactive arthritis; diabetes; multiple sclerosis;  
KM Hashimoto's thyroiditis; Grave's disease; contact dermatitis; psoriasis;  
KM graft rejection; graft versus host disease; actima; glomerulonephritis;  
XX allergy; gene therapy; TH cell subpopulation; ss.  
OS  
XX Mus musculus.  
FH Key location/Qualifiers  
FT CDS 40..885  
FT /\*tag= a  
FT /product= "gene 200 product"  
XX  
PN US6455685-B1.  
XX  
PD 24-SEP-2002.  
XX  
PF 27-FEB-1998; 9805-0032337.  
XX  
PR 03-MAR-1995; 9505-0398633.  
PR 07-JUN-1995; 9505-0487748.  
PR 01-MAR-1996; 9605-0609583.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX PA  
XX PI  
XX Levinson DA;  
XX  
XX MPI; 2003-066247/06.  
XX  
XX P-PSDB; ABG73716.  
XX  
XX Identifying a test compound that binds to a 103 gene product or is a  
XX TH2 cell marker, useful for the identification and therapeutic use of  
XX PT compounds as treatments of helminthic, bacterial and viral infections,  
XX PT and immune disorders -  
XX  
XX Example 8; Figure 17A-D; 131pp; English.  
XX  
XX This invention describes a novel method for identifying a test compound

CC that binds to a 103 gene product or is a TH2 cell marker. The method  
CC comprises contacting a test compound with an immobilised 103 gene  
CC product, removing unbound test compound or separating the complex from  
CC the reaction mixture, and detecting the complex. Identifying a test  
CC compound that binds to a 103 gene product alternatively comprises: (a)  
CC contacting a test compound with a cell engineered to express a 103 gene  
CC product or co-expressing a 103 gene product and a test compound in a  
CC cell, and (b) removing unbound 103 gene product. Identifying a test  
CC compound that is a TH2 cell marker further comprises detecting a complex,  
CC where contacting the test compound with a TH2 and TH1 cell for the test  
CC compound to bind either cell and detecting binding where it indicates the  
CC test compound as a TH2 cell marker. The products described in the  
CC invention have protozoicidal, antibacterial, vitruicide, immunosuppressive,  
CC antiinflammatory, antirheumatic, antidiabetic, neuroprotective,  
CC dermatological, antichryoid, antiporiatic, nephrotoxic, antiaesthetic;  
CC and antiallergic activity and can act as CD8, CD4 and interleukin  
CC agonists. The methods and compositions of the present invention are  
CC useful for the identification and therapeutic use of compounds as  
CC treatments of helminthic, bacterial and viral infections and immune  
CC disorders such as Crohn's disease, reactive arthritis, diabetes,  
CC multiple sclerosis, Hashimoto's thyroiditis, Grave's disease, contact  
CC dermatitis, psoriasis, graft rejection, graft versus host disease,  
CC asthma, allergy and glomerulonephritis. They can also be used for gene  
CC therapy, for the diagnostic evaluation and prognosis of TH cell  
CC subpopulation-related disorders, identification of subjects exhibiting a  
CC predisposition to such conditions, monitoring undergoing clinical  
CC evaluation and efficacy for the treatment of the disorders. This sequence  
CC represents a T-Helper cell associated polynucleotide described in the  
CC disclosure of the invention.

SO Sequence 2710 BP; 789 A; 593 C; 643 G; 678 T; 7 other;

Query Match 99.8%; Score 2705; DB 25; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2708; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGACCCACGGGCTGCGGATTCCTCCCAAGTACTATGTTTTCAGTCTTACCTCA 61  
DB 2 GTGACCCACGGGCTGCGGATTCCTCCCAAGTACTATGTTTTCAGTCTTACCTCA 61  
QY 62 ACTGTCTCTGCTGTGCTGCACTACTTGTCAAGTCAATGTGAAGTGTATTAAG 121  
DB 62 ACTGTCTCTGCTGTGCTGCACTACTTGTCAAGTCAATGTGAAGTGTATTAAG 121  
QY 122 TTGAGGTTGTATTAAGTGTCTTATCTGCTGCACTTCACTTCAATCTGGACAC 181  
DB 122 TTGAGGTTGTATTAAGTGTCTTATCTGCTGCACTTCACTTCAATCTGGACAC 181  
QY 182 TTGTGCTATGTGTGCGGGCAAGGGATTCCTGCTTGTGTCACAGTGTACCAATGAGTTGC 241  
DB 182 TTGTGCTATGTGTGCGGGCAAGGGATTCCTGCTTGTGTCACAGTGTACCAATGAGTTGC 241  
QY 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGCAGATACCAAGGAGCG 301  
DB 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGCAGATACCAAGGAGCG 301  
QY 302 ATCTCAACAAAGAGATGTGTCTTGATATTAAGATGTGACTGTGATGACATGGGA 361  
DB 302 ATCTCAACAAAGAGATGTGTCTTGATATTAAGATGTGACTGTGATGACATGGGA 361  
QY 362 CCTACTGTGACGAGATACAGTTCCTGCTTATGATGATTAATAATTAAGACTGAAT 421  
DB 362 CCTACTGTGACGAGATACAGTTCCTGCTTATGATGATTAATAATTAAGACTGAAT 421  
QY 422 TAGACATCAAGACAGCAAGTCACTCCAGCTCAGACTGCGGAGCTTACTACAG 481  
DB 422 TAGACATCAAGACAGCAAGTCACTCCAGCTCAGACTGCGGAGCTTACTACAG 481  
QY 482 CTTCTCAAGAACCTTACCAAGAGAAATGTTCAGAGACACAGACTGTGTACCC 541  
DB 482 CTTCTCAAGAACCTTACCAAGAGAAATGTTCAGAGACACAGACTGTGTACCC 541  
QY 542 TCATATTAATGAATGAACAAAATTTCCATGAGGCTGATGAATTAAGACTGTGAG 601

|||||  
Db 542 TCCATATATACATGGAACAAAAATTTCCACATGGCTGATGAAATTAAGAGCTCGAG 601  
602 AAACGATCGAAGCTGATCCACATTTGAGTGGAGTCTCTGCTGGTTGACCCCTGGAC 661  
Db 602 AAACGATCGAAGCTGATCCACATTTGAGTGGAGTCTCTGCTGGTTGACCCCTGGAC 661  
662 TTATCATTTGCTGCTTAAATCTTAAATGATATTCCTGTAGAAAAAGAAAGTATGAGTT 721  
Db 662 TTATCATTTGCTGCTTAAATCTTAAATGATATTCCTGTAGAAAAAGAAAGTATGAGTT 721  
722 TGAGCTTATTTACACTGCGCAACTTGCCTCCAGAGGGTTGCAATGACAGAGCATCA 781  
Db 722 TGAGCTTATTTACACTGCGCAACTTGCCTCCAGAGGGTTGCAATGACAGAGCATCA 781  
782 GGATTCGCTCTGAGAAAAATATCTACACATCGAGAGAAAGTATATGAGTGGAGAAAT 841  
Db 782 GGATTCGCTCTGAGAAAAATATCTACACATCGAGAGAAAGTATATGAGTGGAGAAAT 841  
842 CAATGAGTACTACTGCTACGCTCAACAGCGCAGAGCATCTGACCGCTCTGAGCTGCC 901  
Db 842 CAATGAGTACTACTGCTACGCTCAACAGCGCAGAGCATCTGACCGCTCTGAGCTGCC 901  
902 ACTTTAAAGGCTCGCTTCAATTTCTGACTTTGTATTTCCCTTTXTGAAAAATATGTG 961  
Db 902 ACTTTAAAGGCTCGCTTCAATTTCTGACTTTGTATTTCCCTTTXTGAAAAATATGTG 961  
962 AATATGCACTTGGCAACCTCATTGAGGTTCTGACACAGCCCATGAGAAAGAGTTCCA 1021  
Db 962 AATATGCACTTGGCAACCTCATTGAGGTTCTGACACAGCCCATGAGAAAGAGTTCCA 1021  
1022 GTTTTCTGGGGATTAATTACTACAGAGGGATTCGACTGTAATCTAGTACATTGAAT 1081  
Db 1022 GTTTTCTGGGGATTAATTACTACAGAGGGATTCGACTGTAATCTAGTACATTGAAT 1081  
1082 GCTCCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCATCCAGAGCTTCAATCATG 1141  
Db 1082 GCTCCATTTTATCCCTGAGTTTCAGGGATCGGATCTCCCATCCAGAGCTTCAATCATG 1141  
1142 CGTGTGAAAGCTCACTGCTGCTTTCATACATTAAGAAAGTATGCTGATCTTTGAGA 1201  
Db 1142 CGTGTGAAAGCTCACTGCTGCTTTCATACATTAAGAAAGTATGCTGATCTTTGAGA 1201  
1202 CATAGAGTTTGTGATATCCGCAAGCTCCGAAAGGTTGAGGGGAAATTAAGGCTTAA 1261  
Db 1202 CATAGAGTTTGTGATATCCGCAAGCTCCGAAAGGTTGAGGGGAAATTAAGGCTTAA 1261  
1262 GATAGGAAGTCCGCTCTTGTGATGTTGAAAAATCTTAAGAAAGTGTAGCTTTTCT 1321  
Db 1262 GATAGGAAGTCCGCTCTTGTGATGTTGAAAAATCTTAAGAAAGTGTAGCTTTTCT 1321  
1322 AAGAGATTTCTGACTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT 1381  
Db 1322 AAGAGATTTCTGACTTGAAGATTAAGAAAAAGCCAGGTGGCATATGCTTAAACAGATAT 1381  
1382 AACTTGGGAACCTTAGGCAAGAGGGGATTAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441  
Db 1382 AACTTGGGAACCTTAGGCAAGAGGGGATTAAGTTCAAGGTCAGCCAGGCTATGCTGTA 1441  
1442 AGACTGTCTCAGCATCCAAAGAGCAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
Db 1442 AGACTGTCTCAGCATCCAAAGAGCAAAATTAACATAGAGACAGAGAGGCTGAGATGA 1501  
1502 GGCTCGGACAGTGAAGGTGATGTTGTAAGCAAGCAAGAGATCTATATTTGATGTAAGAC 1561  
Db 1502 GGCTCGGACAGTGAAGGTGATGTTGTAAGCAAGCAAGAGATCTATATTTGATGTAAGAC 1561  
1562 CCAAGATGAAAAAGCTTAGGCTGCTAGAGCATGCTTGTAGCTCAGAGATGGAAGGTAA 1621  
Db 1562 CCAAGATGAAAAAGCTTAGGCTGCTAGAGCATGCTTGTAGCTCAGAGATGGAAGGTAA 1621  
1622 AGGCAACAGATCCCGGGGCTTGGCTGACAGCTTAGCTTAGCTGAGTTCCAA 1681  
|||||

Db 1622 AGGCAACAGATCCCGGGGCTTGGCTGACAGCTTAGCTTAGCTGAGTTCCAA 1681  
1682 GTCCACAGAGTCCCTGCTCAGAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741  
Db 1682 GTCCACAGAGTCCCTGCTCAGAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1741  
1742 GGTGTCTCTCTCTCAGAGAGACATGACATGCTGCTGACACACACACACACACACAC 1801  
Db 1742 GGTGTCTCTCTCTCAGAGAGACATGACATGCTGCTGACACACACACACACACACAC 1801  
1802 ACAACACACACACACACACACACACACATGATGAAAGTTCTCTGCTGCTGCTAC 1861  
Db 1802 ACAACACACACACACACACACACACACATGATGAAAGTTCTCTGCTGCTGCTAC 1861  
1862 TCTCTATTAACATGATCTCTCTCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
Db 1862 TCTCTATTAACATGATCTCTCTCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
1922 CATGCGACAGAGCTCAGTAATTTTATTCACAGACTCAGAAAGGCTGAGCAGAGCTGGA 1981  
Db 1922 CATGCGACAGAGCTCAGTAATTTTATTCACAGACTCAGAAAGGCTGAGCAGAGCTGGA 1981  
1982 GAGTTCAAGAGCACTGTGCCCAACACTGCGAGACTCTTCTTAACAGAAAAAGTTACC 2041  
Db 1982 GAGTTCAAGAGCACTGTGCCCAACACTGCGAGACTCTTCTTAACAGAAAAAGTTACC 2041  
2042 CGAAGAGCCCTGCTCTGTAAAGAAAGCCCTGCGAAAGCACTTGACTGTGTG 2101  
Db 2042 CGAAGAGCCCTGCTCTGTAAAGAAAGCCCTGCGAAAGCACTTGACTGTGTG 2101  
2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161  
Db 2102 TGCTCAAGGGGAACTGACTCAGACAACTTCTCATTTCTGAGAGAACTGAGCTGTTTC 2161  
2162 TGACAGAAACCAACCGGTGACTGGGACATACGAAGGAGAGCTTTCAGCAATCTATA 2221  
Db 2162 TGACAGAAACCAACCGGTGACTGGGACATACGAAGGAGAGCTTTCAGCAATCTATA 2221  
2222 TAGTACGAAAAATATCTTTTGGAGAGACAGTCTGTACCAATTTGATTTCCAAAGCCGCTG 2281  
Db 2222 TAGTACGAAAAATATCTTTTGGAGAGACAGTCTGTACCAATTTGATTTCCAAAGCCGCTG 2281  
2282 ACCTCAGTTTATCTGGCTTACAGCTGCTGCCAGTCCCTTGTATCTGTGCTGCC 2341  
Db 2282 ACCTCAGTTTATCTGGCTTACAGCTGCTGCCAGTCCCTTGTATCTGTGCTGCC 2341  
2342 ATCTATAACAGAAATTAATATAGACCCGAGTGAAATATTAAGTGAAGCAAGAGT 2401  
Db 2342 ATCTATAACAGAAATTAATATAGACCCGAGTGAAATATTAAGTGAAGCAAGAGT 2401  
2402 AGCTTTGTTCAAAAGTTTTTTTGCATTTGGGAGCACTGTGTACATCAGAGCACTGT 2461  
Db 2402 AGCTTTGTTCAAAAGTTTTTTTGCATTTGGGAGCACTGTGTACATCAGAGCACTGT 2461  
2462 TAGTGAAGCAACAAACCTGTGATACCGTTTTTTCATGTATGAAATTTTGTGTTAGT 2521  
Db 2462 TAGTGAAGCAACAAACCTGTGATACCGTTTTTTCATGTATGAAATTTTGTGTTAGT 2521  
2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581  
Db 2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581  
2582 TTAACAAATCCATTGAGATTAACAGCTCTCATGCAAGAGGAAACCTAATCTCAATGTT 2641  
Db 2582 TTAACAAATCCATTGAGATTAACAGCTCTCATGCAAGAGGAAACCTAATCTCAATGTT 2641  
2642 TTAAGTATATAAACTGACTGCGCAAGTACTTGAAGATTAATAAAAAAAAAAAAAAAG 2701  
Db 2642 TTAAGTATATAAACTGACTGCGCAAGTACTTGAAGATTAATAAAAAAAAAAAAAAAG 2701  
2702 GCGGCGCGC 2710  
Db 2702 GCGGCGCGC 2710



QY 722 TGAGCTTATTAACATGCGCAACTTGGCTCCAGAGAGGTTGGCAAAATGACAGAGAGTCA 781  
DB 722 TGAGCTTATTAACATGCGCAACTTGGCTCCAGAGAGGTTGGCAAAATGACAGAGAGTCA 781  
QY 782 GGAATTCGCTCTGAGAGAAAATATCTACACCATCGAGAGAACGTATATGAAATGAGAAAT 841  
DB 782 GGAATTCGCTCTGAGAGAAAATATCTACACCATCGAGAGAACGTATATGAAATGAGAAAT 841  
QY 842 CAAAAGAGATCTACGCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901  
DB 842 CAAAAGAGATCTACGCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901  
QY 902 ACTTTTAAAGGCTCGGCTTCACTTCTGACTTGTATTTCCCTTTTGTGAAAATCTATGTC 961  
DB 902 ACTTTTAAAGGCTCGGCTTCACTTCTGACTTGTATTTCCCTTTTGTGAAAATCTATGTC 961  
QY 962 ATATGCTCACTTGGCAACCTCATTTGAGAGTTCTGACCAAGAGAGAGAGAGAGAGAGAG 1021  
DB 962 ATATGCTCACTTGGCAACCTCATTTGAGAGTTCTGACCAAGAGAGAGAGAGAGAGAGAG 1021  
QY 1022 GTTTTCTGAGAGAAATTAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
DB 1022 GTTTTCTGAGAGAAATTAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
QY 1082 GCTCCATTTTATCCCTGAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141  
DB 1082 GCTCCATTTTATCCCTGAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141  
QY 1142 CGTGTGGAAGCTCACTCGTCTTCTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201  
DB 1142 CGTGTGGAAGCTCACTCGTCTTCTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201  
QY 1202 CATAGAGTTTGTGATATCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261  
DB 1202 CATAGAGTTTGTGATATCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261  
QY 1262 GATAGAGAGTGGGCTTCTTGTGATGTTGAGAAATCTTAAAGAGTTGATGCTTTC 1321  
DB 1262 GATAGAGAGTGGGCTTCTTGTGATGTTGAGAAATCTTAAAGAGTTGATGCTTTC 1321  
QY 1321 T-AGAGATTTCTGACCTTGAAGAGATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379  
DB 1322 TAAAGAGATTTCTGACCTTGAAGAGATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1381  
QY 1380 ATTAATCTTGGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439  
DB 1382 ATTAATCTTGGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441  
QY 1440 TAAAGATCTCTGACCTTGAAGAGATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499  
DB 1442 TAAAGATCTCTGACCTTGAAGAGATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501  
QY 1500 GAGGCTCGAGAGAGTGGATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559  
DB 1502 GAGGCTCGAGAGAGTGGATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561  
QY 1560 CCCCAGATGAAAAGCTGAG 1619  
DB 1562 CCCCAGATGAAAAGCTGAG 1621  
QY 1620 AAAAG 1679  
DB 1622 AAAAG 1681  
QY 1680 AAGTCAACAAGAGTCCCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739  
DB 1682 AAGTCAACAAGAGTCCCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741  
QY 1740 GGGGTTGCTCTCTCTGAG 1799  
DB 1742 GGGGTTGCTCTCTCTGAG 1801  
QY 1800 ACACAGACACACACACACACACACACACACATGATGAAAGTTCTCTCTGCTGCTGCTA 1859

DB 1802 ACACAGACACACACACACACACACACACACATGATGAAAGTTCTCTCTGCTGCTGCTA 1861  
QY 1860 CCTCTCTTAACATATATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1919  
DB 1862 CCTCTCTTAACATATATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1921  
QY 1920 AGCATGAG 1979  
DB 1922 AGCATGAG 1981  
QY 1980 GAGAGTTGAG 2039  
DB 1982 GAGAGTTGAG 2041  
QY 2040 CCGGAG 2099  
DB 2042 CCGGAG 2101  
QY 2100 TGTGCTCAAGGGGAG 2159  
DB 2102 TGTGCTCAAGGGGAG 2161  
QY 2160 TCTGACAG 2219  
DB 2162 TCTGACAG 2221  
QY 2220 TATAGTCAAG 2279  
DB 2222 TATAGTCAAG 2281  
QY 2280 GGAAGCTCAAGTTTCACTGAG 2339  
DB 2282 GGAAGCTCAAGTTTCACTGAG 2341  
QY 2340 CCATCTATTAACAG 2399  
DB 2342 CCATCTATTAACAG 2401  
QY 2400 GTAGCTTTGTTAAAGATTTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2459  
DB 2402 GTAGCTTTGTTAAAGATTTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461  
QY 2460 GTTAG 2519  
DB 2462 GTTAG 2521  
QY 2520 GTTGTCTTCTAGCTGAG 2579  
DB 2522 GTTGTCTTCTAGCTGAG 2581  
QY 2580 TCTTAACAAATCAATTAAG 2639  
DB 2582 TCTTAACAAATCAATTAAG 2641  
QY 2640 TTTTAAAGTAAATCAATTAAG 2699  
DB 2642 TTTTAAAGTAAATCAATTAAG 2701  
QY 2700 AGGGGGGGGGC 2710  
DB 2702 AGGGGGGGGGC 2712  
RESULT 9  
ABZ68328  
ABZ68328 standard; DNA; 2725 BP.  
XX  
AC ABZ68328;  
XX  
DT 22-APR-2003 (first entry)  
XX  
DB Nucleotide sequence of murine TIM-3 BALB/c allele.

XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
KW allergic T cell response; autoimmune disease; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS /tag= a  
FT /product= "TIM-3"  
PN MO2003002722-A2.  
XX  
PD 09-JUN-2003.  
XX  
PF 01-JUL-2002; 2002MO-US20890.  
XX  
PR 29-JUN-2001; 2001US-302344P.  
XX  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;  
XX  
DR MPI: 2003-210268/20.  
XX  
P-PSDB; ABR70434.  
XX  
PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
XX  
PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
XX  
PT allergy, eczema or autoimmune disease -  
XX  
PS Claim 5; Page 76; 94pp; English.

XX The present sequence encodes a murine T cell immunoglobulin domain and  
XX mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
XX TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
XX conserved IgV and mucin domains. The locus comprising the TIM family is  
XX genetically associated with immune dysfunction, including asthma. The  
XX TIM gene family is located within a region of human chromosome 5 that  
XX is commonly deleted in malignancies and myelodysplastic syndrome.  
XX Variants of TIM-1 and TIM-3 are associated with susceptibility to  
XX airway hyperreactivity and allergic T cell responses, and other variants  
XX associated with protection against these responses. T cells express TIM  
XX proteins, which critically regulate CD4 T cell differentiation. Th1  
XX cells preferentially express TIM-3, while Th2 cells preferentially  
XX express TIM-1. TIM polypeptides and polynucleotides are useful for  
XX treating cancer, asthma, allergies, eczema or autoimmune diseases.

XX Sequence 2725 BP; 805 A; 587 C; 641 G; 692 T; 0 other;

Query Match 96.8%; Score 2624.6; DB 25; Length 2725;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2655; Conservative 6; Mismatches 20; Indels 2; Gaps 2;

QY 18 GGAATTCCTCCCTCCAGTACTGATGTTTCAAGTCTTACCTCAAGTGTCTGCTGCT 77  
DB 45 GGAATTCCTCCCTCCAGTACTGATGTTTCAAGTCTTACCTCAAGTGTCTGCTGCT 104  
QY 78 GCTGCAACTACTACTGTCGCAAGTTCATGGAAGTGTATTAAGTTGAGTTGGTAAAAA 137  
DB 105 GCTGCAACTACTACTGTCGCAAGTTCATGGAAGTGTATTAAGTTGAGTTGGTAAAAA 164  
QY 138 TGCCTATTCGCTCGAGTTCACCTTACCTATCATCTGGGACCTTGTGCTATGTGCTG 197  
DB 165 TGCCTATTCGCTCGAGTTCACCTTACCTATCATCTGGGACCTTGTGCTATGTGCTG 224  
QY 198 GGGCAAGGAGATTCGCTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAG 257  
DB 225 GGGCAAGGAGATTCGCTGTCAGAGTGTACCAATGAGTGTCTGAGAACTGATGAAG 284  
QY 258 AATGTGACATATCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAACAAAGAGA 317  
DB 285 AATGTGACATATCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAACAAAGAGA 344  
QY 318 TGTGTCCTGATCATTAAGATGTGACTCTGATGATCCATGGGACCTACTGCTGAGAT 377  
DB 345 TGTGTCCTGATCATTAAGATGTGACTCTGATGATCCATGGGACCTACTGCTGAGAT 404  
QY 378 ACAGTTCCTGCTTATGAATGAATTAAGAACTGAAGTAAATTAAGATCAAGACAC 437  
DB 405 ACAGTTCCTGCTTATGAATGAATTAAGAACTGAAGTAAATTAAGATCAAGACAC 464  
QY 438 CAAGTCACTCCAGTCAAGTCAAGTCCATGAGGAGCTTACTAGAGCTTCTCAAGACCT 497  
DB 465 CAAGTCACTCCAGTCAAGTCAAGTCCATGAGGAGCTTACTAGAGCTTCTCAAGACCT 524  
QY 498 AACCAAGGAGAAATGTTTCAGAGACACAGACACTGCTGCTCATTAATPAACATGG 557  
DB 525 AACCAAGGAGAAATGTTTCAGAGACACAGACACTGCTGCTCATTAATPAACATGG 584  
QY 558 AACCAAAATTTCCACATGGGCTGATGAATTAAGGACTCTGGAGAAAGATGAGACTGC 617  
DB 585 AACCAAAATTTCCACATGGGCTGATGAATTAAGGACTCTGGAGAAAGATGAGACTGC 644  
QY 618 TATCCACATTTGAGTGGAGTCTGCTGGGTTGACCTTGACCTTATCATTTGCTGCT 677  
DB 645 TATCCACATTTGAGTGGAGTCTGCTGGGTTGACCTTGACCTTATCATTTGCTGCT 704  
QY 678 AATCTTAAATGTTATCTCTGTAAGAAAAGATTAATGAGTTTGAAGCTTATTAACCT 737  
DB 705 AATCTTAAATGTTATCTCTGTAAGAAAAGATTAATGAGTTTGAAGCTTATTAACCT 764  
QY 738 GGCCAACTGCTCCAGAGGAGTTGGCAAAATCAGAGACATCAGATTCGCTGAGAGA 797  
DB 765 GGCCAACTGCTCCAGAGGAGTTGGCAAAATCAGAGACATCAGATTCGCTGAGAGA 824  
QY 798 AATATCTACACATCAGAGAGAGATATATGAGTGAATGAGATTAAGTACTACTG 857  
DB 825 AATATCTACACATCAGAGAGAGATATATGAGTGAATTAAGTACTACTG 884  
QY 858 TAACTCAACAGCCAGCAGCCATCTGACCGCTCTGACCTGCACTTTTAAAGCTGCG 917  
DB 885 TAACTCAACAGCCAGCAGCCATCTGACCGCTCTGACCTGCACTTTTAAAGCTGCG 944  
QY 918 CTTCAATTTCTGCTTGGATTTCCCTTTTGGAAAACATATGATGATGCTGGGCA 977  
DB 945 CTTCAATTTCTGCTTGGATTTCCCTTTTGGAAAACATATGATGATGCTGGGCA 1003  
QY 978 CCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAGATTCAGTTTCTGGGATAT 1037  
DB 1004 CCTCATTTGAGGTTCTGACCAAGCCACTGAGAAAAGATTCAGTTTCTGGGATAT 1063  
QY 1038 TAACTCAACAGGAGATTTGAGCTGTAATCTATGATTAATGATGATTTATTCCT 1097  
DB 1064 TAACTCAACAGGAGATTTGAGCTGTAATCTATGATTAATGATGATTTATTCCT 1123  
QY 1098 GAGTTTCAGGATTCGATCTCCCACTCCAGAGACTTCAATCATGCTGTTGAAGCTACT 1157  
DB 1124 GAGTTTCAGGATTCGATCTCCCACTCCAGAGACTTCAATCATGCTGTTGAAGCTACT 1183  
QY 1158 CGTGTCTTCAATTAAGAAATGTTAGTGTATCTTTGAGACATAGAGGTTTGTGT 1217  
DB 1184 CGTGTCTTCAATTAAGAAATGTTAGTGTATCTTTGAGACATAGAGGTTTGTGT 1243  
QY 1218 AATATCCGAAAGCTCTGAAACAGTGAAGGGAATTAAGGCTTAATATGAAAGTGGCT 1277  
DB 1244 AATATCCGAAAGCTCTGAAACAGTGAAGGGAATTAAGGCTTAATATGAAAGTGGCT 1303  
QY 1278 CTTTGTGATGTTGAAATCTTAAGAAGTGTGAGCTTTCTGAGATTTCTGACCT 1337  
DB 1304 TCTTTGTGATGTTGAAATCTTAAGAAGTGTGAGCTTTCTGAGATTTCTGACCT 1362  
QY 1338 GAAAGATTAAGAAAAGCCAGGTGCTATATGCTTAACAGATTAATCTTGGAACTTAG 1397  
DB 1363 GAAAGATTAAGAAAAGCCAGGTGCTATATGCTTAACAGATTAATCTTGGAACTTAG 1422

QY 1398 GCAGAGGGGTATAGTTCAAGTTCAGCCAGGGCTATGCTGGTAGACCTGCTCCAMCMTTC 1457  
 DB 1423 GCAGAGGGGTATAGTTCAAGTTCAGCCAGGGCTATGCTGGTAGACCTGCTCCAMCMTTC 1482  
 QY 1458 CAAGAAGCAAAATTAACCTAGAGACAGAGAGGCTGGAGATGAGCTGGACAGTGAAGG 1517  
 DB 1483 CAAGAAGCAAAATTAACCTAGAGACAGAGAGGCTGGAGATGAGCTGGACAGTGAAGG 1542  
 QY 1518 TGCATTGTGTAACAGCAGAGGAATCTATATTGATGCTAGAGCCCATATGAAAAAGCTA 1577  
 DB 1543 TGCATTGTGTAACAGCAGAGGAATCTATATTGATGCTAGAGCCCATATGAAAAAGCTA 1602  
 QY 1578 GGCCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAAAGGACAAACAGATCCC 1637  
 DB 1603 GGCCTGTAGAGCATGCTTGTAGACTCAAGAGATGAGAGGTAAAGGACAAACAGATCCC 1662  
 QY 1638 CGGGGCTTGGCTGACGTAGCTTAGCCTTAGAGTGTGAGTTCAGAGTCCAGAGTCCCT 1697  
 DB 1663 CGGGGCTTGGCTGACGTAGCTTAGCCTTAGAGTGTGAGTTCAGAGTCCAGAGTCCCT 1722  
 QY 1698 GTCTCAGAGTAAGATGAGCTGATCTGGGCGATGCTCATGGGGGTGCTCCTCCTCT 1757  
 DB 1723 GTCTCAGAGTAAGATGAGCTGATCTGGGCGATGCTCATGGGGGTGCTCCTCCTCT 1782  
 QY 1758 CAGAAGACATGACATGACCTGACACACACACACACACACACACACACACACACACAC 1817  
 DB 1783 CAGAAGACATGACATGACCTGACACACACACACACACACACACACACACACACACAC 1842  
 QY 1818 ACACACACACATGATAATGAAGGTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1877  
 DB 1843 ACACACACACATGATAATGAAGGTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902  
 QY 1878 CTCTCAGAGTAAGATGAGCTGATCTGGGCGATGCTCATGGGGGTGCTCCTCCTCT 1937  
 DB 1903 CTCTCAGAGTAAGATGAGCTGATCTGGGCGATGCTCATGGGGGTGCTCCTCCTCT 1962  
 QY 1938 AGTAATTTATTCACACACTCAGAGGCTGAGACAGAGGCTGAGAGTTCAGAGACACTG 1997  
 DB 1963 AGTAATTTATTCACACACTCAGAGGCTGAGACAGAGGCTGAGAGTTCAGAGACACTG 2022  
 QY 1998 TGCCCAACACCTGCGACACTCTTCTTACACAGAAAAAGTTACCGGCAAGAGCTGCTG 2057  
 DB 2023 TGCCCAACACCTGCGACACTCTTCTTACACAGAAAAAGTTACCGGCAAGAGCTGCTG 2082  
 QY 2058 TCTGTAAAAGAAACCTGCGAAAGCAAACTTTGACTGTGTGTGCTCAAGGGAACTG 2117  
 DB 2083 TCTGTAAAAGAAACCTGCGAAAGCAAACTTTGACTGTGTGTGCTCAAGGGAACTG 2142  
 QY 2118 ACTCAGACAACTTCTCATTTCTGAGAGAACTGAGAGTGTTCAGACAGAAACAAC 2177  
 DB 2143 ACTCAGACAACTTCTCATTTCTGAGAGAACTGAGAGTGTTCAGACAGAAACAAC 2202  
 QY 2178 GGTGACTGGGACATACGAAGGAGAGCTTTCAGCAATCTATATAGTCAGCAAAATAT 2237  
 DB 2203 GGTGACTGGGACATACGAAGGAGAGCTTTCAGCAATCTATATAGTCAGCAAAATAT 2262  
 QY 2238 CTTTGGGAGGAGATGCTGCAACCAATGATTTCCAAAGCCGGTGGACCTTCAGTTTCACTG 2297  
 DB 2263 CTTTGGGAGGAGATGCTGCAACCAATGATTTCCAAAGCCGGTGGACCTTCAGTTTCACTG 2322  
 QY 2298 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2357  
 DB 2323 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2382  
 QY 2358 AATTAATTAAGACCCGAGTGAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2417  
 DB 2383 AATTAATTAAGACCCGAGTGAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2442  
 QY 2418 TTTTTCGATTTGGGAGCACTGTGTACATCAGAGACATCTGTATAGTAGAGACACCAAA 2477  
 DB 2443 TTTTTCGATTTGGGAGCACTGTGTACATCAGAGACATCTGTATAGTAGAGACACCAAA 2502

QY 2478 ACTGTGTGATCCGTTTTTTCATGATGATGATTTGTTAGTTGCTTACCTAGCTAGCTG 2537  
 DB 2503 ACTGTGTGATCCGTTTTTTCATGATGATGATTTGTTAGTTGCTTACCTAGCTAGCTG 2562  
 QY 2538 TGGAGGTCCTGGCTTTCTTACGTGGGATGAGAGGACCATCTAACAAATCCATTAG 2597  
 DB 2563 TGGAGGTCCTGGCTTTCTTACGTGGGATGAGAGGACCATCTAACAAATCCATTAG 2622  
 QY 2598 AGATPACAGCTCTCATGCGAAGGAAAACTATCTCAATGTTTTTAATTAATTAAT 2657  
 DB 2623 AGATPACAGCTCTCATGCGAAGGAAAACTATCTCAATGTTTTTAATTAATTAAT 2682  
 QY 2658 GTACTGGCAAGTACTTTGACCATMAAAAAAAAAAAAAA 2700  
 DB 2683 GTACTGGCAAGTACTTTGACCATMAAAAAAAAAAAAAA 2725  
 RESULT 10  
 AB268329  
 ID AB268329 standard; DNA; 862 BP.  
 AC AB268329;  
 DT 22-APR-2003 (first entry)  
 XX  
 XX  
 XX  
 DT Nucleotide sequence of murine TIM-3 ES-HBA and DBA/2J allele.  
 DE  
 XX  
 XX  
 KM T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;  
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;  
 KM myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;  
 KM allergic T cell response; autoimmune disease; gene; ss.  
 OS  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT CDS 17..862  
 FT /\*tag= a  
 FT /product= "TIM-3"  
 PN W02003002722-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 PF 01-JUL-2002; 2002WO-US20890.  
 XX  
 PR 29-JUN-2001; 2001US-302344P.  
 XX  
 PA (STRD ) UNIV LEIAND STANFORD JUNIOR.  
 XX  
 PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GD, Kuchroo V;  
 DR WPI, 2003-210268/20.  
 DR F-PSDB; ABP70435.  
 XX  
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain  
 PT and Mucin domain gene sequences, useful for treating cancer or asthma,  
 PT allergy, eczema or autoimmune disease -  
 XX  
 PS Claim 5, Page 77; 94pp; English.  
 CC The present sequence encodes a murine T cell immunoglobulin domain and  
 CC mucin domain-3 (TIM-3) polypeptide. The specification describes TIM-1,  
 CC TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with  
 CC conserved Igv and mucin domains. The locus comprising the TIM family is  
 CC genetically associated with immune dysfunction, including asthma. The  
 CC TIM gene family is located within a region of human chromosome 5 that  
 CC is commonly deleted in malignancies and myelodysplastic syndrome.  
 CC Variants of TIM-1 and TIM-3 are associated with susceptibility to  
 CC airway hyperreactivity and allergic T cell responses, and other variants  
 CC associated with protection against these responses. T cells express TIM  
 CC proteins, which critically regulate CD4 T cell differentiation. Th1  
 CC cells preferentially express TIM-3, while Th2 cells preferentially  
 CC express TIM-1. TIM polypeptides and polynucleotides are useful for







QY	340	GGAGCTCTGGAGTACCATGATGGAGCTTACCTGTGAGAGTACAGTTCCCTGGTCTTAAAGAT	399
Db	301	GTGACTCTGGATGACCATGGAGCTTACTCTGTGAGATACAGTTCCCTGGTCTTAAAGAT	360
QY	400	GATAAAAAATTAGAACTGAAATTAAGACATCAAGCAGCCAGGTCACTCAGCTCAGACT	459
Db	361	GATAAAAAATTAGAACTGAAATTAAGACATCAAGCAGCCAGGTCACTCAGCTCAGACT	420
QY	460	GCCCATGGGGACTCTTACTACAGTTCTCCAGAACCTTAAACGAGAGAAATGGTTCA	519
Db	421	GCCCATGGGGACTCTTACTACAGTTCTCCAGAACCTTAAACGAGAGAAATGGTTCA	480
QY	520	GAGACACAGACACTGCTGACCTCCATAATTAACAATGSAACAAAATTTTCCATGGGGCT	579
Db	481	GAGACACAGACACTGCTGACCTCCATAATTAACAATGSAACAAAATTTTCCATGGGGCT	540
QY	580	GATGAAATTTAAGACTCTGTGAGAAACGATCAGATCGTATCCACATTTGGAGTGGAGTTC	639
Db	541	GATGAAATTTAAGACTCTGTGAGAAACGATCAGATCGTATCCACATTTGGAGTGGAGTTC	600
QY	640	TCTGCTGGGTTGACCTTGCACTTATCATTGGTGTCTTAATCCTTAATGTAATTCCTGT	699
Db	601	TCTGCTGGGTTGACCTTGCACTTATCATTGGTGTCTTAATCCTTAATGTAATTCCTGT	660
QY	700	AAGAAAAAGAAATTTACAGTTTGAGCCTTAATACACTGGCCAACTTGCTCCAGAGGG	759
Db	661	AAGAAAAAGAAATTTACAGTTTGAGCCTTAATTAATACACTGGCCAACTTGCTCCAGAGGG	720
QY	760	TTGGCAAAATGCAGAGACGTACAGATTCGCTCTGAGAGAAATATATCTACACATCGAGAG	819
Db	721	TTGGCAAAATGCAGAGACGTACAGATTCGCTCTGAGAGAAATATATCTACACATCGAGAG	780
QY	820	AACGTATATGAATGAGAGAAATCAAAATGAGTACTACTGCTAACAGCCAGCAGCCCA	879
Db	781	AACGTATATGAATGAGAGAAATCAAAATGAGTACTACTGCTAACAGCCAGCAGCCCA	840
QY	880	TCC 882	
Db	841	TCC 843	

RESULT 12	
AA170255	
ID	AA170255 standard; cDNA, 843 BP.
XX	
AC	AA170255;
XX	
DT	07-JAN-2002 (first entry)
DE	Mouse 200 gene, preferentially expressed in TH1 cells.
XX	
KW	Mouse; 200 gene; T helper; T lymphocyte; T cell; TH1;
KM	receptor; differential expression, immune disorder; psoriasis; multiple sclerosis; insulin-dependent diabetes; antidiabetic; antipsoriatic; diagnosis; therapy; ss.
XX	
OS	Mus musculus.
XX	
PN	US6288218-B1.
XX	
PD	11-SEP-2001.
XX	
PF	25-SEP-1997; 97US-0937399.
XX	
PR	01-MAR-1996; 96US-0609583.
PR	03-MAR-1995; 95US-0398633.
XX	
PA	07-JUN-1995; 95US-0487748.
XX	
PI	(LEVI) LEVINSON D A.
XX	
PI	Levinson DA;
DR	WPI; 2001-647189/74.

DR	P-PSDB; AAM50225.
XX	
PT	Detecting 200 gene expression in a sample, useful for treating and
PT	diagnosing immune disorders, especially T lymphocyte-related disorders
PT	comprises detecting the presence of a 200 gene product or an RNA
PT	encoding the 200 gene product -
XX	
PS	Disclosure: Column 123-124; 108pp; English.

The present sequence is that of cDNA comprising the complete open reading frame of the mouse 200 gene. It encodes a 281-amino acid protein (see AM50225). The cDNA was isolated using the 'T cell clone paradigm' and the 'transgenic T cell paradigm' approaches to the identification of 'genes that are differentially expressed in helper T cell (TH) subpopulations'. The gene is expressed at a higher level in fully differentiated TH1 than in TH2 cells. It can be used diagnostically or as a target for therapeutic intervention for the treatment of immune disorders. A claimed method for diagnosing a TH cell subpopulation-related immune disorder involves detecting the level of a human 200 gene (see AA10234) product, or an RNA encoding it, so that if the level differs from that in a control sample, the disorder is diagnosed. The disorder is especially a TH1 cell subpopulation-related immune disorder, such as multiple sclerosis, psoriasis or insulin-dependent diabetes (claimed). A method for detecting murine 200 gene expression is also claimed. In addition to the 200 gene, the invention provides other genes that are differentially expressed within and among TH cells and TH cell subpopulations and which can be used in methods for the diagnosis, prognosis, evaluation and treatment of TH cell subpopulation-related disorders, for the identification of subjects exhibiting a predisposition to such conditions, for monitoring patients undergoing clinical evaluation for the treatment of such disorders, and for monitoring the efficacy of compounds used in clinical trials. Other immune disorders that can be treated/diagnosed include Crohn's disease, reactive arthritis, Lyme disease, organ-specific autoimmunity, Hashimoto's thyroiditis, Grave's disease, contact dermatitis, graft rejection, graft versus host disease, sarcoidosis, atopic conditions, asthma, allergy, allergic rhinitis, food allergy, eosinophilia, conjunctivitis, glomerular nephritis, helminthic infection (e.g. leishmaniasis), viral infection (e.g. HIV), and bacterial infection (e.g. tuberculosis and lepromatous leprosy).

Sequence 843 BP; 253 A; 185 C; 196 G; 209 T; 0 other;

Query Match	31.1%	Score 843	DB 22	Length 843
-------------	-------	-----------	-------	------------

Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 40 AGCTTTTCAGGCTCTTAACCCCTCAACATGTGTGCCTGGCTGCGAATACTACTCTTGAAGG 99

Dd 1 AAGTTTTAGGCTTAAACCCTCAACATGTGTGCCTGGCTGCGAATACTACTCTTGAAGG 60

QY 100 TCATTGGAAGAATGATTATTAAGGTTGAGCTTGCTGTAATAAATGCCCATCTGCCCCAGTTAC 159

Dd 61 TCAATTGGAAGAATGATTATTAAGGTTGAGCTTGCTGTAATAAATGCCCATCTGCCCCAGTTAC 120

QY 160 ACTCTAACCTACATCTCGGGAACACTTGTGCCATATGTCTGGGGCMAAGGATTTCTGCTTGG 219

Dd 121 ACTCTAACCTACATCTCGGGAACACTTGTGCCATATGTCTGGGGCMAAGGATTTCTGCTTGG 180

QY 220 TCACAGTGTACCAATGATAGTTGCTCAGAACCTGATGAAAGAAATGTGACATATCCAGAAATCC 279

Dd 181 TCACAGTGTACCAATGATAGTTGCTCAGAACCTGATGAAAGAAATGTGACATATCCAGAAATCC 240

QY 280 AACCAATACCACTAAAGGCGATCTCAACAAAGAGATGTGTCTTGATCATAAAGAAAT 339

Dd 241 AACCAATACCACTAAAGGCGATCTCAACAAAGAGATGTGTCTTGATCATAAAGAAAT 300

QY 340 GTGACCTCTGATGAGCACATGTGGGACCTATCTGCTGAGAGATACAGTTCCCTGGCTTTAAGAAT 399

Dd 301 GTGACCTCTGATGAGCACATGTGGGACCTATCTGCTGAGAGATACAGTTCCCTGGCTTTAAGAAT 360

QY 400 GATATAAAATTAGAACTGAAATTAGACATCAAGGACCAAGGCTCACTCCAGCTCAACT 459  
 |||  
 DB 361 GATATAAAATTAGAACTGAAATTAGACATCAAGGACCAAGGCTCACTCCAGCTCAACT 420  
 |||  
 QY 460 GCCCATGGGAGCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAAATGGTTCA 519  
 |||  
 DB 421 GCCCATGGGAGCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAAATGGTTCA 480  
 |||  
 QY 520 GAGACACAGACACTGGTGACCCCTCCATAATTAACATGAAACAAAATTTCCACATGGGCT 579  
 |||  
 DB 481 GAGACACAGACACTGGTGACCCCTCCATAATTAACATGAAACAAAATTTCCACATGGGCT 540  
 |||  
 QY 580 GATGAATTAAGACTCTGGAGAAACGATCAAGACCTGCTATCCACATTTGGAGTGGAGTC 639  
 |||  
 DB 541 GATGAATTAAGACTCTGGAGAAACGATCAAGACCTGCTATCCACATTTGGAGTGGAGTC 600  
 |||  
 QY 640 TCTGCTGGTGAACCTCGGACCTTATGCTGTCTTAATCCCTTAATGATGATTCCTGT 659  
 |||  
 DB 601 TCTGCTGGTGAACCTCGGACCTTATGCTGTCTTAATCCCTTAATGATGATTCCTGT 660  
 |||  
 QY 700 AAGAAAAAGAACTTATGAGTTTGAGCTTATTAACCTGGCCAACTTGCCTCCAGAGGG 759  
 |||  
 DB 661 AAGAAAAAGAACTTATGAGTTTGAGCTTATTAACCTGGCCAACTTGCCTCCAGAGGG 720  
 |||  
 QY 760 TTGGCAAAATGAGAGAGCTCAGGATTCGCTCTGAGGAAATATCTACACCATCGAGAG 819  
 |||  
 DB 721 TTGGCAAAATGAGAGAGCTCAGGATTCGCTCTGAGGAAATATCTACACCATCGAGAG 780  
 |||  
 QY 820 AACGTATATGAAAGTGAAGAAATTCAAATGAGTACTACTGCTACGTCACAGCCAGAGCCA 879  
 |||  
 DB 781 AACGTATATGAAAGTGAAGAAATTCAAATGAGTACTACTGCTACGTCACAGCCAGAGCCA 840  
 |||  
 QY 880 TCC 882  
 |||  
 DB 841 TCC 843  
 |||

RESULT 13  
 AAF82615  
 ID AAF82615 standard; cDNA; 843 BP.  
 AC AAF82615;  
 XX  
 DT 18-JUN-2001 (first entry)  
 XX  
 DE Murine TH1 specific 200 gene coding sequence.  
 XX  
 KM Murine; T helper cell; TH cell; TH1; TH2; immunomodulator;  
 KM anti-inflammatory; antiallergic; dermatological; antiviral;  
 KM antibacterial; T helper lymphocyte modulator; gene therapy;  
 KM TH specific gene; 200 gene; immune disorder; inflammation;  
 KM infection; ss.  
 XX  
 OS Mus sp.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..843  
 FT /tag= a  
 FT /product= "200 gene product"  
 XX  
 PN US6204371-B1.  
 XX  
 PD 20-MAR-2001.  
 XX  
 PF 01-MAR-1996; 96US-0609583.  
 XX  
 PR 03-MAR-1995; 95US-0398633.  
 PR 07-JUN-1995; 95US-0487748.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Levinson DA;  
 XX

DR WPI; 2001-272703/28.  
 XX  
 XX New murine or human 200 genes and their corresponding polypeptides,  
 PT useful for treating or diagnosing immune disorders, especially T helper  
 PT lymphocyte-related disorders, e.g. inflammatory diseases or allergies  
 PT  
 PS Claim 6; Fig 17; 109pp; English.  
 XX  
 CC The present sequence is expressed at higher levels in T helper 1 (TH1)  
 CC cells than in T helper 2 (TH2) cells. The invention relates to an  
 CC isolated nucleic acid molecule, which comprises the full length murine  
 CC 200 gene or full length human 200 gene nucleotide sequence. The nucleic  
 CC acids are useful for treating or diagnosing immune disorders,  
 CC especially T helper lymphocyte-related disorders, e.g. inflammatory  
 CC diseases (e.g. Crohn's disease), multiple sclerosis, Grave's disease,  
 CC contact dermatitis, psoriasis, asthma and allergies, or certain viral  
 CC (e.g. HIV) or bacterial (e.g. tuberculosis) infections.  
 XX  
 SQ Sequence 843 BP; 253 A; 185 C; 196 G; 209 T; 0 other;  
 Query Match 31.1%; Score 843; DB 22; Length 843;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-227;  
 Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 40 ATGTTTTCAGGTTTACCTCTCAACTGTGCTCTGCTGCTGCACTACTACTCTTCAAG 99  
 |||  
 DB 1 ATGTTTTCAGGTTTACCTCTCAACTGTGCTCTGCTGCTGCTGCACTACTACTCTTCAAG 60  
 |||  
 QY 100 TCATTGGAAGATGTTATAGAGTTGAGTTGGTGAATGAATGCTATCTGCTCCAGTTAC 159  
 |||  
 DB 61 TCATTGGAAGATGTTATAGAGTTGAGTTGGTGAATGAATGCTATCTGCTCCAGTTAC 120  
 |||  
 QY 160 ACTTACCTTAACCTTGGACACTTGTGCTATGTGCTGAGGCAAGGATTTGCTTGG 219  
 |||  
 DB 121 ACTTACCTTAACCTTGGACACTTGTGCTATGTGCTGAGGCAAGGATTTGCTTGG 180  
 |||  
 QY 220 TCACAGTGAACCAATGATGTTGCTCAGAACTGATGAAGAAATGTGACATTCAGAAATCC 279  
 |||  
 DB 181 TCACAGTGAACCAATGATGTTGCTCAGAACTGATGAAGAAATGTGACATTCAGAAATCC 240  
 |||  
 QY 280 AGCAGATACAGCTTAAGGGGCGATCTCAACAAGGAGATGTGCTCTGATCAATTAAGAT 339  
 |||  
 DB 241 AGCAGATACAGCTTAAGGGGCGATCTCAACAAGGAGATGTGCTCTGATCAATTAAGAT 300  
 |||  
 QY 340 GTGACTCTGATGACCATGAGGAGCTTACTGCTGAGATACAGTTCCCTGCTTATGAAT 399  
 |||  
 DB 301 GTGACTCTGATGACCATGAGGAGCTTACTGCTGAGATACAGTTCCCTGCTTATGAAT 360  
 |||  
 QY 400 GATATAAAATTGAATGAATGAATTAAGCATCAAGAGCCCAAGGTCTCCAGCTCAGACT 459  
 |||  
 DB 361 GATATAAAATTGAATGAATGAATTAAGCATCAAGAGCCCAAGGTCTCCAGCTCAGACT 420  
 |||  
 QY 460 GCCCATGGGAGCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAAATGGTTCA 519  
 |||  
 DB 421 GCCCATGGGAGCTCTACTACAGCTTCTCAAGAACCTTAACACGAGAGAAATGGTTCA 480  
 |||  
 QY 520 GAGACACAGACACTGGTGACCCCTCCATAATTAACATGAAACAAAATTTCCACATGGGCT 579  
 |||  
 DB 481 GAGACACAGACACTGGTGACCCCTCCATAATTAACATGAAACAAAATTTCCACATGGGCT 540  
 |||  
 QY 580 GATGAATTAAGACTCTGGAGAAACGATCAAGACCTGCTATCCACATTTGGAGTGGAGTC 639  
 |||  
 DB 541 GATGAATTAAGACTCTGGAGAAACGATCAAGACCTGCTATCCACATTTGGAGTGGAGTC 600  
 |||  
 QY 640 TCTGCTGGTGAACCTCGGACCTTATGCTGTCTTAATCCCTTAATGATGATTCCTGT 659  
 |||  
 DB 601 TCTGCTGGTGAACCTCGGACCTTATGCTGTCTTAATCCCTTAATGATGATTCCTGT 660  
 |||  
 QY 700 AAGAAAAAGAACTTATGAGTTTGAGCTTATTAACCTGGCCAACTTGCCTCCAGAGGG 759  
 |||  
 DB 661 AAGAAAAAGAACTTATGAGTTTGAGCTTATTAACCTGGCCAACTTGCCTCCAGAGGG 720  
 |||





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 21, 2003, 16:48:30 ; Search time 166.515 Seconds  
(without alignments)  
7183.411 Million cell updates/sec

Title: US-10-004-633-8

Perfect score: 2710

Sequence: 1 ngctgaccacgcgtccgca.....aaaaaaaaagcgccgcgc 2710

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2706.6	99.9	2710	1	US-08-487-748A-8
2	2706.6	99.9	2710	3	US-08-480-070C-8
4	2706.6	99.9	2710	3	US-08-829-525-8
5	2706.6	99.9	2710	3	US-08-609-583A-8
6	2706.6	99.9	2710	3	US-08-937-399-8
7	2706.6	99.9	2710	4	US-09-310-367-8
8	2706.6	99.9	2710	4	US-09-032-337-8
9	843	31.1	843	3	US-08-829-525-22
10	843	31.1	843	3	US-08-609-583A-22
11	843	31.1	843	3	US-08-937-399-22
12	843	31.1	843	4	US-09-310-367-22
13	843	31.1	843	4	US-09-032-337-22
14	843	31.1	843	4	US-09-464-231-22
15	499.8	18.4	2236	3	US-08-829-525-23
16	499.8	18.4	2236	3	US-08-609-583A-23
17	499.8	18.4	2236	3	US-08-937-399-23
18	499.8	18.4	2236	4	US-09-310-367-23
19	499.8	18.4	2236	4	US-09-032-337-23
20	499.8	18.4	2236	4	US-09-464-231-23
21	456	16.8	903	3	US-08-829-525-37
22	456	16.8	903	3	US-08-609-583A-37
23	456	16.8	903	3	US-08-937-399-37
24	456	16.8	903	4	US-09-310-367-37
25	456	16.8	903	4	US-09-032-337-37
26	456	16.8	903	4	US-09-464-231-37
27	240.8	8.9	313	3	US-08-398-633-8

28	64.4	2.4	1380	4	US-09-620-312D-59	Sequence 59, Appl
29	62.8	2.3	7218	1	US-08-232-463-14	Sequence 14, Appl
30	62	2.3	2093	1	US-08-287-001A-1	Sequence 1, Appl
31	62	2.3	2093	5	PCT-US95-09941-1	Sequence 1, Appl
32	60.8	2.2	2310	3	US-08-458-434A-1	Sequence 1, Appl
33	60.8	2.2	9299	3	US-08-458-434A-7	Sequence 7, Appl
34	60.4	2.2	80595	3	US-09-078-294-3	Sequence 3, Appl
35	59	2.2	538	4	US-09-495-050A-105	Sequence 105, App
36	58.8	2.2	80246	3	US-09-078-294-4	Sequence 4, Appl
37	57.6	2.1	3358	3	US-08-957-063-4	Sequence 4, Appl
38	57.6	2.1	3358	4	US-09-487-685-4	Sequence 4, Appl
39	57.6	2.1	3358	4	US-08-802-805D-4	Sequence 4, Appl
40	56	2.1	3381	3	US-09-009-119-1	Sequence 1, Appl
41	56	2.1	3381	4	US-09-371-507-1	Sequence 1, Appl
42	56	2.1	3383	5	PCT-US95-09098-1	Sequence 1, Appl
43	55.4	2.0	46718	4	US-09-816-093-3	Sequence 3, Appl
44	55.2	2.0	200	4	US-09-354-147C-31	Sequence 31, Appl
45	55.2	2.0	564	1	US-08-117-362-32	Sequence 32, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-487-748A-8
; Sequence 8, Application US/08487748A
; Patent No. 5721351
;
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487, 748A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO. 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 40..885
; US-08-487-748A-8
;
; Query Match 99.9%; Score 2706.6; DB 1; Length 2710;
; Best Local Similarity 100.0%; Pred. No. 0;
; Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 2 GTCGACCAACGCGTCGATTTCCCTCCCAAGTACTGTTTCAAGCTTACCCATCA 61
```

Db 2 GTGACCCACGCGTCCGGATTTCCCTCCCAAGTACTCATGTTTACAGGCTTTACCTTCA 61  
Qy 62 ACTGTGCTGTGCTGTGCTGCACTA CTATTGCAAGGTCAATTGGAAGTGTATTAAG 121  
Db 62 ACTGTGCTGTGCTGTGCTGCACTA CTATTGCAAGGTCAATTGGAAGTGTATTAAG 121  
Qy 122 TTGAAGTTGGTAAAAATGCTATCTGCTGCGAGTTACACTTACCTTACATCTGGACAC 181  
Db 122 TTGAAGTTGGTAAAAATGCTATCTGCTGCGAGTTACACTTACCTTACATCTGGACAC 181  
Qy 182 TTGTGCTATGTGCTGGGCAAGGATCTGTCTTGTGTCACAGTGCACCAATGAGTTG 241  
Db 182 TTGTGCTATGTGCTGGGCAAGGATCTGTCTTGTGTCACAGTGCACCAATGAGTTG 241  
Qy 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGCAGATACAGCTAAAGGAG 301  
Db 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATCCAGCAGATACAGCTAAAGGAG 301  
Qy 302 ATCTCAACAAAGAGATGTGTCTGTGATCATAAAGATGTGACTGTGATGACATGGGA 361  
Db 302 ATCTCAACAAAGAGATGTGTCTGTGATCATAAAGATGTGACTGTGATGACATGGGA 361  
Qy 362 CCTACTGCTGCAGATACAGTTCCCTGTGCTTATGATGATTAATAAATTAGAACTGAAT 421  
Db 362 CCTACTGCTGCAGATACAGTTCCCTGTGCTTATGATGATTAATAAATTAGAACTGAAT 421  
Qy 422 TAGACATCAAGAGCAGAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 481  
Db 422 TAGACATCAAGAGCAGAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 481  
Qy 482 CTCTCTCAAGAACCTTACACCGAGAGAAATGTTTCAAGACACAGACATCTGTGACCC 541  
Db 482 CTCTCTCAAGAACCTTACACCGAGAGAAATGTTTCAAGACACAGACATCTGTGACCC 541  
Qy 542 TCCATATATACAAATGAAACAAAATTTCCACATGGGCTGATGAATTAAGAACTGAG 601  
Db 542 TCCATATATACAAATGAAACAAAATTTCCACATGGGCTGATGAATTAAGAACTGAG 601  
Qy 602 AAACGATCAGAACTGCTATCCATTTGAGTGGAGTCTCTGTGGGTGACCCCTGGCAC 661  
Db 602 AAACGATCAGAACTGCTATCCATTTGAGTGGAGTCTCTGTGGGTGACCCCTGGCAC 661  
Qy 662 TTTATCTGTGTCTTAAATCTTAAATGTAATCTGTGTAAGAAAGAAAGTTATCGAGTT 721  
Db 662 TTTATCTGTGTCTTAAATCTTAAATGTAATCTGTGTAAGAAAGAAAGTTATCGAGTT 721  
Qy 722 TGAGCCTTATTAACATGCGCACTTGTGCTCAGAGGAGTTGGCAATGACAGAGCTGA 781  
Db 722 TGAGCCTTATTAACATGCGCACTTGTGCTCAGAGGAGTTGGCAATGACAGAGCTGA 781  
Qy 782 GGAATTCGCTCTGAGAGAAATATCTACACCATCGAGAGAAAGTATATGAAGTGAAGTT 841  
Db 782 GGAATTCGCTCTGAGAGAAATATCTACACCATCGAGAGAAAGTATATGAAGTGAAGTT 841  
Qy 842 CAAATGAGTACTCTGCTAGTCAACAGCAGAGCACTCTGACCGCTCTGAGTCTGC 901  
Db 842 CAAATGAGTACTCTGCTAGTCAACAGCAGAGCACTCTGACCGCTCTGAGTCTGC 901  
Qy 902 ACTTTAAAGGCTCGGCTTCAATTTGACTTGTGATTTTCCCTTTTGTGAAAATCTATGTT 961  
Db 902 ACTTTAAAGGCTCGGCTTCAATTTGACTTGTGATTTTCCCTTTTGTGAAAATCTATGTT 961  
Qy 962 ATATGCTACTTGGCAACTCTATTTGAGGTTCTGACCAAGCCACTGAGAAAAGAGTTTCA 1021  
Db 962 ATATGCTACTTGGCAACTCTATTTGAGGTTCTGACCAAGCCACTGAGAAAAGAGTTTCA 1021  
Qy 1022 GTTTTGTGGGGAATTAATCTCAAGGGGATTCGACTGATCTCACTTGAAT 1081  
Db 1022 GTTTTGTGGGGAATTAATCTCAAGGGGATTCGACTGATCTCACTTGAAT 1081  
Qy 1082 GCTCATTTTATCCCTGAGTTTCAAGGATCGATCTCCCACTCAAGAGCTTCAATCATG 1141  
Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATCGATCTCCCACTCAAGAGCTTCAATCATG 1141

Db 1082 GCTCATTTTATCCCTGAGTTTCAAGGATCGATCTCCCACTCAAGAGCTTCAATCATG 1141  
Qy 1142 CGTGTGAAGCTCACTCGTCTTTTATATCATTTAGAAATGTTAGTGTGATCTTTGAGA 1201  
Db 1142 CGTGTGAAGCTCACTCGTCTTTTATATCATTTAGAAATGTTAGTGTGATCTTTGAGA 1201  
Qy 1202 CATTAAGTTTGTGTATATCCGCAAGCTCCTGAAACAGTGTGGGGAATTAAGGCTTA 1261  
Db 1202 CATTAAGTTTGTGTATATCCGCAAGCTCCTGAAACAGTGTGGGGAATTAAGGCTTA 1261  
Qy 1262 GATTAAGAGTGCAGTCTTTTGTGATGTTGAAAATCTTAAGAAATGTTGATCTTTTCT 1321  
Db 1262 GATTAAGAGTGCAGTCTTTTGTGATGTTGAAAATCTTAAGAAATGTTGATCTTTTCT 1321  
Qy 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAAGTGTGATGCTTAACAGATAT 1381  
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAAGTGTGATGCTTAACAGATAT 1381  
Qy 1382 AACTTGGGAACCTTGAAGAGGAGTGAATGTTCAAGTGCAGAGGCTGATGCTGTA 1441  
Db 1382 AACTTGGGAACCTTGAAGAGGAGTGAATGTTCAAGTGCAGAGGCTGATGCTGTA 1441  
Qy 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATAGAGACAGAGAGCTGAGATGA 1501  
Db 1442 AGACTGTCTCAACATCCAAAAGCAAAATTAACATAGAGACAGAGAGCTGAGATGA 1501  
Qy 1502 GGCCTGGAACATGAGTGTATTTGTACAAAGCAGAGAAATCTATATTTGATCGTAGAC 1561  
Db 1502 GGCCTGGAACATGAGTGTATTTGTACAAAGCAGAGAAATCTATATTTGATCGTAGAC 1561  
Qy 1562 CCAACATGAAGAAAGCTAGGCTGTGAGAGATCTTGTAGACTCAAGAGATGAAGAGTGA 1621  
Db 1562 CCAACATGAAGAAAGCTAGGCTGTGAGAGATCTTGTAGACTCAAGAGATGAAGAGTGA 1621  
Qy 1622 AGGCAACAAGATCCCGGGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1681  
Db 1622 AGGCAACAAGATCCCGGGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1681  
Qy 1682 GTTCAACAAGATCCCTGTCTCAMAAGTAAAGTGTGATCTGAGGATCTGATGAG 1741  
Db 1682 GTTCAACAAGATCCCTGTCTCAMAAGTAAAGTGTGATCTGAGGATCTGATGAG 1741  
Qy 1742 GGTGTGCTCTCTCAGAGAGAGACATGCAATGCTGTGACACACACACACACAC 1801  
Db 1742 GGTGTGCTCTCTCAGAGAGAGACATGCAATGCTGTGACACACACACACACACAC 1801  
Qy 1802 ACACACACACACACACACACACACACACATGATGAAGTCTCTGTGCTGTGAC 1861  
Db 1802 ACACACACACACACACACACACACACACATGATGAAGTCTCTGTGCTGTGAC 1861  
Qy 1862 TCTCTATTAATGATGATCTCTACAGAGCTCTCTGTGCTGTGATGATGAGTGGAG 1921  
Db 1862 TCTCTATTAATGATGATCTCTACAGAGCTCTCTGTGCTGTGATGATGAGTGGAG 1921  
Qy 1922 CATGGCAGAGCAGTCAAGTAAATTTTTCAGAGCTCAGAAAGCTGAGAGAGTGA 1981  
Db 1922 CATGGCAGAGCAGTCAAGTAAATTTTTCAGAGCTCAGAAAGCTGAGAGAGTGA 1981  
Qy 1982 GAGTTCAAGAGCAGTGTGCTCAACACTGCAAGCTCTTCTTACACAGAAAAGTTACC 2041  
Db 1982 GAGTTCAAGAGCAGTGTGCTCAACACTGCAAGCTCTTCTTACACAGAAAAGTTACC 2041  
Qy 2042 CGCAAGCAGCTGTGCTGTAAAGAAACCTGCAAGGCAAACTTTGACTGTTGG 2101  
Db 2042 CGCAAGCAGCTGTGCTGTGTAAAGAAACCTGCAAGGCAAACTTTGACTGTTGG 2101  
Qy 2102 TGTCTAAGGGGAATCTGACACAACTTCTCATTTCTGGAGAAATCTGAGCTGTTTC 2161  
Db 2102 TGTCTAAGGGGAATCTGACACAACTTCTCATTTCTGGAGAAATCTGAGCTGTTTC 2161  
Qy 2162 TGACAGAAAGCAACCGGTGATCTGGGACATAGAAAGGAGAGACTTGTGACGAACTATA 2221  
Db 2162 TGACAGAAAGCAACCGGTGATCTGGGACATAGAAAGGAGAGACTTGTGACGAACTATA 2221



QY 2222 TAGTCAGCAAAATATCTTTGGGAGAGACGTCACCAATGATTTCCAAGCCGGTGG 2281  
DB 2222 TAGTCAGCAAAATATCTTTGGGAGAGACGTCACCAATGATTTCCAAGCCGGTGG 2281  
QY 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGTSCCTTGATCTGTCTGGCTCC 2341  
DB 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGTSCCTTGATCTGTCTGGCTCC 2341  
QY 2342 ATCTATAACAGATCAATTAATTAATGACCCGAGGAAATTTTAAGTACAGCAAAAGT 2401  
DB 2342 ATCTATAACAGATCAATTAATTAATGACCCGAGGAAATTTTAAGTACAGCAAAAGT 2401  
QY 2402 AGCTTTGTTCAAGATTTTGTGTCATTTGGGAGCACTGTGATCATCTAGAGACATCTGT 2461  
DB 2402 AGCTTTGTTCAAGATTTTGTGTCATTTGGGAGCACTGTGATCATCTAGAGACATCTGT 2461  
QY 2462 TAGTGAGGACACCAAACTGTGTGTAACGTTTTCATGTATGAATTTTGTGTTAGT 2521  
DB 2462 TAGTGAGGACACCAAACTGTGTGTAACGTTTTCATGTATGAATTTTGTGTTAGT 2521  
QY 2522 TGCTTTAGCTAGCTGTGAGAGTCTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581  
DB 2522 TGCTTTAGCTAGCTGTGAGAGTCTGCTTTCTTAAGTGGTATGGAAGGAGACATC 2581  
QY 2582 TAACAAAATCATTAGATTAAGATTAACAGCTCTCATGAGAGGAAACTAATCTCAATGTT 2641  
DB 2582 TAACAAAATCATTAGATTAAGATTAACAGCTCTCATGAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTAAGTAAATAAACTGTACTGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
DB 2642 TTAAGTAAATAAACTGTACTGCGCAAGTACTTTGAGCATTAATAAAAAAAAAAAG 2701  
QY 2702 GCGGCGCGC 2710  
DB 2702 GCGGCGCGC 2710

## RESULT 2

US-08-480-070C-8  
Sequence 8, Application US/08480070C  
Patent No. 6066498  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,070C  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-480-070C-8  
Query Match 99.9%; Score 2706.6; DB 3; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GTGACCCAGCGTCGCGGATTTCCCTCCCAAGTACATCTTTCAGTCTTACCCCA 61  
DB 2 GTGACCCAGCGTCGCGGATTTCCCTCCCAAGTACATCTTTCAGTCTTACCCCA 61  
QY 62 ACTGTGTCCTGCTGCTGCACTACTTTCGAAGGTCAATGGAAGATGTTAAG 121  
DB 62 ACTGTGTCCTGCTGCTGCACTACTTTCGAAGGTCAATGGAAGATGTTAAG 121  
QY 122 TTGAGTTGTAAATAATGCTATCTGCTGCACTTCACTTACATCTGAGAC 181  
DB 122 TTGAGTTGTAAATAATGCTATCTGCTGCACTTCACTTACATCTGAGAC 181  
QY 182 TTGAGCTATGCTGTGGGCAAGGATTTCTGCTGTCAGTGTACATATAGTTC 241  
DB 182 TTGAGCTATGCTGTGGGCAAGGATTTCTGCTGTCAGTGTACATATAGTTC 241  
QY 242 TCAGAACTGATGAAGAATAATGTACATATCAGAAATCAGAGATPACGCTAAAGSGCG 301  
DB 242 TCAGAACTGATGAAGAATAATGTACATATCAGAAATCAGAGATPACGCTAAAGSGCG 301  
QY 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTACTGTGATGACCAATGGA 361  
DB 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTACTGTGATGACCAATGGA 361  
QY 362 CCTACTGTGAGATATACGTTCCCTGCTTTAATGAATGAATTAAGATCTGAAT 421  
DB 362 CCTACTGTGAGATATACGTTCCCTGCTTTAATGAATGAATTAAGATCTGAAT 421  
QY 422 TAGACATCAAGCAGCCAGGTCACCTGCTGAGCTGCCATGSGGAGCTCTACTAG 481  
DB 422 TAGACATCAAGCAGCCAGGTCACCTGCTGAGCTGCCATGSGGAGCTCTACTAG 481  
QY 482 CTCTCCAGAACCTTAACACGAGAGAAATGTTTCAAGACACACACTGTGTACC 541  
DB 482 CTCTCCAGAACCTTAACACGAGAGAAATGTTTCAAGACACACACTGTGTACC 541  
QY 542 TCCATTAATACATGGAACAAATTTTCCATGAGGCTGATGAATTAAGATCTGTGAG 601  
DB 542 TCCATTAATACATGGAACAAATTTTCCATGAGGCTGATGAATTAAGATCTGTGAG 601  
QY 602 AAACGATCAGAACTGCTATCCATGGAAGGAGTCTGCTGGGTGACCCCTGGAC 661  
DB 602 AAACGATCAGAACTGCTATCCATGGAAGGAGTCTGCTGGGTGACCCCTGGAC 661  
QY 662 TTATCATTTGTTGCTTTAATCTTAATGTTTCTGTTAAGAAAGAAAGTTATCGAGTT 721  
DB 662 TTATCATTTGTTGCTTTAATCTTAATGTTTCTGTTAAGAAAGAAAGTTATCGAGTT 721  
QY 722 TGAGCTTTATACACTGGCCCACTTGGCTCCAGAGAGGTTGGCAATGACAGAGATCA 781  
DB 722 TGAGCTTTATACACTGGCCCACTTGGCTCCAGAGAGGTTGGCAATGACAGAGATCA 781  
QY 782 GGATTCGCTGAGGAAATATCTACACATCGAGAGAACTATATAGTGAATTT 841  
DB 782 GGATTCGCTGAGGAAATATCTACACATCGAGAGAACTATATAGTGAATTT 841  
QY 842 CAAATGAGTACTACTGTACGTCACACGACGACCATCTGACCGGCTCTGAGCTGCC 901  
DB 842 CAAATGAGTACTACTGTACGTCACACGACGACCATCTGACCGGCTCTGAGCTGCC 901



QY 902 ACTTTAAAGGCTGGCTTCATTTCTGACCTTNGTATTTCCCTTTKGGAAAATCTATG 961  
 DB 902 ACTTTAAAGGCTGGCTTCATTTCTGACCTTNGTATTTCCCTTTKGGAAAATCTATG 961  
 QY 962 ATATGCACTTGGCAACCTCATTTGAGAGTTCTGACCAAGCCACTGAGAAAAGATGCCA 1021  
 DB 962 ATATGCACTTGGCAACCTCATTTGAGAGTTCTGACCAAGCCACTGAGAAAAGATGCCA 1021  
 QY 1022 GTTTTCTGGGGATTAATTAATCAACAGGGATTCGACTGATCTCATCTCAATTTGAAT 1081  
 DB 1022 GTTTTCTGGGGATTAATTAATCAACAGGGATTCGACTGATCTCATCTCAATTTGAAT 1081  
 QY 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATGAGATCCCACTCCAGAGATTTCAATCAG 1141  
 DB 1082 GCTCCATTTTATCCCTGAGTTTCAAGGATGAGATCCCACTCCAGAGATTTCAATCAG 1141  
 QY 1142 CGTGTGAAGCTCACTGTCCTTTCATACATTAAGAAATGTTAGTGTCTTTGAGA 1201  
 DB 1142 CGTGTGAAGCTCACTGTCCTTTCATACATTAAGAAATGTTAGTGTCTTTGAGA 1201  
 QY 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAAACAGTAGGGGAAATAAAGGCTAA 1261  
 DB 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAAACAGTAGGGGAAATAAAGGCTAA 1261  
 QY 1262 GATAGGAAGGTGGGTCTTGTGATGTTGGAATAATCTTAAGAAGTGTGACTTTCT 1321  
 DB 1262 GATAGGAAGGTGGGTCTTGTGATGTTGGAATAATCTTAAGAAGTGTGACTTTCT 1321  
 QY 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCCAGGTGGCATATGCTTAACAGATAT 1381  
 DB 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCCAGGTGGCATATGCTTAACAGATAT 1381  
 QY 1382 AACTTGGGAACCTTAGGCAAGAGGTGATTAAGTTCAAGGTCAAGCCAGGGCTATGCTGTA 1441  
 DB 1382 AACTTGGGAACCTTAGGCAAGAGGTGATTAAGTTCAAGGTCAAGCCAGGGCTATGCTGTA 1441  
 QY 1442 AGACTGTCTCAGATCCAAAGAGAAATTAACATGAGAGACACAGAGAGCTGAGATGA 1501  
 DB 1442 AGACTGTCTCAGATCCAAAGAGAAATTAACATGAGAGACACAGAGAGCTGAGATGA 1501  
 QY 1502 GGCTCGGACAGTGAAGGTGATGTTGTAACAGCAGAGATCTATATTTGATGCTAGACC 1561  
 DB 1502 GGCTCGGACAGTGAAGGTGATGTTGTAACAGCAGAGATCTATATTTGATGCTAGACC 1561  
 QY 1562 CCACATGAAAAGCTTAGGCTGTGAGCATGCTTGTAGACTCAAGAGATGAGAGATA 1621  
 DB 1562 CCACATGAAAAGCTTAGGCTGTGAGCATGCTTGTAGACTCAAGAGATGAGAGATA 1621  
 QY 1622 AGGCAACAAGATCCCGGGGCTTGGTGAGTCACTTGAAGCTTGAAGTGTGCTTCA 1681  
 DB 1622 AGGCAACAAGATCCCGGGGCTTGGTGAGTCACTTGAAGCTTGAAGTGTGCTTCA 1681  
 QY 1682 GTTCCAAAGATCCCTGTCTCAMAAGTGAAGTGAATGCTGAGTATGCGCATGCGG 1741  
 DB 1682 GTTCCAAAGATCCCTGTCTCAMAAGTGAAGTGAATGCTGAGTATGCGCATGCGG 1741  
 QY 1742 GGTTCCTCTCTCTCTCAGAGAGATGCAATGACCTTGCACACACACACACACAC 1801  
 DB 1742 GGTTCCTCTCTCTCTCAGAGAGATGCAATGACCTTGCACACACACACACACAC 1801  
 QY 1802 AC 1861  
 DB 1802 AC 1861  
 QY 1862 TCTCTAATCATGTATCTCTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 DB 1862 TCTCTAATCATGTATCTCTACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1921  
 QY 1922 CATGGCAGAGCATGCTCAATTAATTTATTCAGACCTCAAGAGCTGAGACAGAGCTGGA 1981  
 DB 1922 CATGGCAGAGCATGCTCAATTAATTTATTCAGACCTCAAGAGCTGAGACAGAGCTGGA 1981

QY 1982 GAGTTCAGAGGACATGTCGCCCAACCTGCCAGACTCTTCTTACACAGAAAAGGTACC 2041  
 DB 1982 GAGTTCAGAGGACATGTCGCCCAACCTGCCAGACTCTTCTTACACAGAAAAGGTACC 2041  
 QY 2042 CCAGAGCAGCTGCTGTCTGTAAAAAGAAACCCTGCAAGAGCAATCTTGACTGTG 2101  
 DB 2042 CCAGAGCAGCTGCTGTCTGTAAAAAGAAACCCTGCAAGAGCAATCTTGACTGTG 2101  
 QY 2102 TGTCTAAGGGGAATCTGACTCAGACAACTTCTCATTTCTGAGAAATCTGAGCTGTT 2161  
 DB 2102 TGTCTAAGGGGAATCTGACTCAGACAACTTCTCATTTCTGAGAAATCTGAGCTGTT 2161  
 QY 2162 TGACAGAGAAACAACCGGTGACTGAGACATACAGAGGAGAGACTCTTGACAGAACTATA 2221  
 DB 2162 TGACAGAGAAACAACCGGTGACTGAGACATACAGAGGAGAGACTCTTGACAGAACTATA 2221  
 QY 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTGACCAAAATTTGCAAGCCGTTG 2281  
 DB 2222 TAGTCAGCAAAATATTTCTTGGAGAGACAGTGTGACCAAAATTTGCAAGCCGTTG 2281  
 QY 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCCAGTGTCTTGAATCTGTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCCAGTGTCTTGAATCTGTGCTGCC 2341  
 QY 2342 ATCTAATACAGAAATCAATTAATTAAGACCCGAGTGAATAATTAAGTGAAGAGAAAGT 2401  
 DB 2342 ATCTAATACAGAAATCAATTAATTAAGACCCGAGTGAATAATTAAGTGAAGAGAAAGT 2401  
 QY 2402 AGCTTGTTCAAAGATTTTCTTGGAGAGCACTGTGTACATCAGAGGACATCTGT 2461  
 DB 2402 AGCTTGTTCAAAGATTTTCTTGGAGAGCACTGTGTACATCAGAGGACATCTGT 2461  
 QY 2462 TAGTGAGACACCAAAACCTGTGTACAGCTTTTTCATGTATGAATTTGTTAGGT 2521  
 DB 2462 TAGTGAGACACCAAAACCTGTGTACAGCTTTTTCATGTATGAATTTGTTAGGT 2521  
 QY 2522 TGCTTCTAGCTAGCTGTGAGGTCCGCTTCTTAAAGTGTATGAAGAGGAGACATC 2581  
 DB 2522 TGCTTCTAGCTAGCTGTGAGGTCCGCTTCTTAAAGTGTATGAAGAGGAGACATC 2581  
 QY 2582 TAACAAAATCATTAGAGATTAACAGCTCTCATGACAGAAAGGAAAATTAATCTCAATGTT 2641  
 DB 2582 TAACAAAATCATTAGAGATTAACAGCTCTCATGACAGAAAGGAAAATTAATCTCAATGTT 2641  
 QY 2642 TTAAGTAAATTAATCTGACTGCAAGTCTTGAAGCTTAATAAAAAAAAAAAAAAG 2701  
 DB 2642 TTAAGTAAATTAATCTGACTGCAAGTCTTGAAGCTTAATAAAAAAAAAAAAAAG 2701  
 QY 2702 GGGGGGCGC 2710  
 DB 2702 GGGGGGCGC 2710

RESULT 3  
 US-08-829-525-8  
 ; Sequence 8, Application US/08829525  
 ; Patent No. 6084083  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
 ; NUMBER OF SEQUENCES: 38  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036/2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/829,525  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-081  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-829-525-8

Query Match 99.9%; Score 2706.6; DB 3; Length 2710;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCACCCACGCGTCGGATTTCCCTCCCAAGTACTGATTTTCAGGTCTTACCTTCA 61  
DB 2 GTCACCCACGCGTCGGATTTCCCTCCCAAGTACTGATTTTCAGGTCTTACCTTCA 61  
QY 62 ACTGTGTCTGTGCTGCTGCACTACTGCAAGTCAATTTGAAGTGTAAAG 121  
DB 62 ACTGTGTCTGTGCTGCTGCACTACTGCAAGTCAATTTGAAGTGTAAAG 121  
QY 122 TTGAGGTTGTAAATGCTATCTGCTGCACTACTGCAAGTCAATTTGAAG 181  
DB 122 TTGAGGTTGTAAATGCTATCTGCTGCACTACTGCAAGTCAATTTGAAG 181  
QY 182 TTGCTCTATGTGCTGGGCAAGGATTTCTGCTGCTGCAAGTCAATTTGAAG 241  
DB 182 TTGCTCTATGTGCTGGGCAAGGATTTCTGCTGCTGCAAGTCAATTTGAAG 241  
QY 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATTCAGAGATCAAGCTAAAGGCG 301  
DB 242 TCAGAACTGATGAAGAAATGTGACATATCAGAAATTCAGAGATCAAGCTAAAGGCG 301  
QY 302 ATCTCAAGAAAGGAGATGTCTCTGATCAATAAGATGTGCTGATGACATGGA 361  
DB 302 ATCTCAAGAAAGGAGATGTCTCTGATCAATAAGATGTGCTGATGACATGGA 361  
QY 362 CCTACTGTGCAAGTACAGTTCCTGTGCTTATGATTAAGTAAATTTAGAACTGAAT 421  
DB 362 CCTACTGTGCAAGTACAGTTCCTGTGCTTATGATTAAGTAAATTTAGAACTGAAT 421  
QY 422 TAGACATCAAGAGCCCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 481  
DB 422 TAGACATCAAGAGCCCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 481  
QY 482 CTTTCCAAAGACCTTAACACGAGAGAAATGTTGAGAGACAGACACGAGTGAAG 541  
DB 482 CTTTCCAAAGACCTTAACACGAGAGAAATGTTGAGAGACAGACACGAGTGAAG 541  
QY 542 TCCATTAATTAACAAATGAAACAAAATTTTCCACATGGGCTGATGATAATTAAGACTGAG 601

DB 542 TCCATTAATTAACAAATGAAACAAAATTTTCCACATGGGCTGATGATAATTAAGACTGAG 601  
QY 602 AAAGATCAAGACCTGATTCACATTTGAGAGGAGTCTGCTGGTTGACCTCGGAC 661  
DB 602 AAAGATCAAGACCTGATTCACATTTGAGAGGAGTCTGCTGGTTGACCTCGGAC 661  
QY 662 TTATCATTTGCTGCTTAATCTTAATGATGATTTCTGATTAAGAAAGAAAGTATGAT 721  
DB 662 TTATCATTTGCTGCTTAATCTTAATGATGATTTCTGATTAAGAAAGAAAGTATGAT 721  
QY 722 TGAGCTTATTAACATGAGCAATGCTGCTGCAAGAGGTTGGCAATGCAAGAGCTCA 781  
DB 722 TGAGCTTATTAACATGAGCAATGCTGCTGCAAGAGGTTGGCAATGCAAGAGCTCA 781  
QY 782 GGATTTGCTGAGAGAAATTTTCAACATTCAGAGAGAAAGTATTAAGTGAAGATT 841  
DB 782 GGATTTGCTGAGAGAAATTTTCAACATTCAGAGAGAAAGTATTAAGTGAAGATT 841  
QY 842 CAAATGATTAATGCTGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901  
DB 842 CAAATGATTAATGCTGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901  
QY 902 ACTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTTCCCTTTTGGAAAGATGATG 961  
DB 902 ACTTTAAAGGCTGCTTCAATTTCTGACTTTGATTTTCCCTTTTGGAAAGATGATG 961  
QY 962 ATATGCTACTGGCAACCTCATTTGAGAGGTTCTGACACAGAGAGAGAGAGAGAG 1021  
DB 962 ATATGCTACTGGCAACCTCATTTGAGAGGTTCTGACACAGAGAGAGAGAGAGAG 1021  
QY 1022 GTTTTCTGGGAGATTAATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
DB 1022 GTTTTCTGGGAGATTAATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081  
QY 1082 GCTCCATTTTATCCCTGATTTCAAGAGATTCGATTCCTCCACTCCAGAGACTTCAATG 1141  
DB 1082 GCTCCATTTTATCCCTGATTTCAAGAGATTCGATTCCTCCACTCCAGAGACTTCAATG 1141  
QY 1142 CGTGTGAAGCTCACTGCTTTTCAATTAAGAAATGATGATGATGATGATGATG 1201  
DB 1142 CGTGTGAAGCTCACTGCTTTTCAATTAAGAAATGATGATGATGATGATGATG 1201  
QY 1202 CATAGAGTTGTGATATCGCAAGAGCTGCTGCAAGAGAGAGAGAGAGAGAGAG 1261  
DB 1202 CATAGAGTTGTGATATCGCAAGAGCTGCTGCAAGAGAGAGAGAGAGAGAGAG 1261  
QY 1262 GATAGAGAGGCTGCTTTTGTGATGATGATGATGATGATGATGATGATGATG 1321  
DB 1262 GATAGAGAGGCTGCTTTTGTGATGATGATGATGATGATGATGATGATGATG 1321  
QY 1322 AGAGATTTCTGATCCTTGAAGAAATTAAGAAAGAGAGAGAGAGAGAGAGAGAT 1381  
DB 1322 AGAGATTTCTGATCCTTGAAGAAATTAAGAAAGAGAGAGAGAGAGAGAGAT 1381  
QY 1382 AACTTGGGAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441  
DB 1382 AACTTGGGAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441  
QY 1442 AGAGCTGCTCAACATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501  
DB 1442 AGAGCTGCTCAACATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501  
QY 1502 GCTTGGAG 1561  
DB 1502 GCTTGGAG 1561  
QY 1562 CCACATGAAG 1621  
DB 1562 CCACATGAAG 1621  
QY 1622 AGGCAACAGATCCCGGAGCTTGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAG 1681

Db 1622 AGGACACAGATCCCGGGGCTTGCGTGCAGTCACTTAGCTAGTGTCCAA 1681  
Qy 1682 GTCCACAGAGTCCCTGTCTCAGAGTGAAGTGTGAGTATCTGGGCGATGTCATGGG 1741  
Db 1682 GTCCACAGAGTCCCTGTCTCAGAGTGAAGTGTGAGTATCTGGGCGATGTCATGGG 1741  
Qy 1742 GGTGTCTCTCTCTCAGAGAGATGACATGACCCCTGACACACACACACACACAC 1801  
Db 1742 GGTGTCTCTCTCTCAGAGAGATGACATGACCCCTGACACACACACACACACAC 1801  
Qy 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATG 1861  
Db 1802 ACACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATG 1861  
Qy 1862 TCTCTATACATGATCTCTACAGAGTCTCTCTGCTCTGTTAAGACATGATGAGAG 1921  
Db 1862 TCTCTATACATGATCTCTACAGAGTCTCTCTGCTCTGTTAAGACATGATGAGAG 1921  
Qy 1922 CATGGAGAGCAGTCCAGTAAATTTATTCAGACCTGAGAGGCTGAGAGAGCGTTGGA 1981  
Db 1922 CATGGAGAGCAGTCCAGTAAATTTATTCAGACCTGAGAGGCTGAGAGAGCGTTGGA 1981  
Qy 1982 GAGTTCAGAGAGCAGTCTGCCCCAACCTGCGAGACTCTTTACACAGAGAGAGTTACC 2041  
Db 1982 GAGTTCAGAGAGCAGTCTGCCCCAACCTGCGAGACTCTTTACACAGAGAGAGTTACC 2041  
Qy 2042 CGCAGAGAGCCTGCTGTCTGTAAAGAGAACCTGCGAGAGAGAGAGAGTTGTTG 2101  
Db 2042 CGCAGAGAGCCTGCTGTGTAAAGAGAACCTGCGAGAGAGAGAGAGTTGTTG 2101  
Qy 2102 TGCTCAAGGGGAACTGACTCAGACCACTTCTGAGAGAGAGAGAGAGAGAGAGAG 2161  
Db 2102 TGCTCAAGGGGAACTGACTCAGACCACTTCTGAGAGAGAGAGAGAGAGAGAGAG 2161  
Qy 2162 TGAAG 2221  
Db 2162 TGAAG 2221  
Qy 2222 TAGTCAGCAAAATATTTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2281  
Db 2222 TAGTCAGCAAAATATTTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2281  
Qy 2282 ACCCTGAGTTTCACTGGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2341  
Db 2282 ACCCTGAGTTTCACTGGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2341  
Qy 2342 ATCTATACAGAGATCAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2401  
Db 2342 ATCTATACAGAGATCAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2401  
Qy 2402 AGCTTTGTTCAAAATTTTTCATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461  
Db 2402 AGCTTTGTTCAAAATTTTTCATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461  
Qy 2462 TAGTGAAG 2521  
Db 2462 TAGTGAAG 2521  
Qy 2522 TGCTTCTAGTCTGAG 2581  
Db 2522 TGCTTCTAGTCTGAG 2581  
Qy 2582 TAAAG 2641  
Db 2582 TAAAG 2641  
Qy 2642 TTAAG 2701  
Db 2642 TTAAG 2701  
Qy 2702 GGGGCGCGC 2710  
Db 2702 GGGGCGCGC 2710

RESULT 4  
US-08-609-583A-8  
Sequence 8, Application US/08609583A  
Patent No. 6204371  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8664  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-609-583A-8  
Query Match 99.9%; Score 2706.6; DB 3; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 GTGACCCACGCGTCCGAGTTTCCCTCCCAAGTACTATGTTTCAAGTCTTACCTCA 61  
Db 2 GTGACCCACGCGTCCGAGTTTCCCTCCCAAGTACTATGTTTCAAGTCTTACCTCA 61  
Qy 62 ACTGTCCTGCTGCTGCTGCACTACTTGAAGTCAATGGAAGTGTATTAAG 121  
Db 62 ACTGTCCTGCTGCTGCTGCACTACTTGAAGTCAATGGAAGTGTATTAAG 121  
Qy 122 TTGAGTTGTAATAATGCTATCTGCTCCGAGTTTACCTTACATCTGGAGAC 181  
Db 122 TTGAGTTGTAATAATGCTATCTGCTCCGAGTTTACCTTACATCTGGAGAC 181  
Qy 182 TTGTCCTATGTCCTGGGCAAGGATTTCTGTCCTTGGTCAAGTGTACCAATGAATTG 241  
Db 182 TTGTCCTATGTCCTGGGCAAGGATTTCTGTCCTTGGTCAAGTGTACCAATGAATTG 241  
Qy 242 TCAGACATGATGAAGAAATGTGACATATCAGAAATCCAGACAGATACAGTTAAAGGGG 301

Db 242 TCAGAACTGATGAAAGAAATGACATATCAGAAATCCAGACGATACCAAGCTAAAGGCG 301  
Qy 302 ATTCGAAGAAAGGAGATGCTCTGATCATATAAAGTGACTCTGATGACCATGGGA 361  
Db 302 ATCTCAAGAAAGGAGATGCTCTGATCATATAAAGTGACTCTGATGACCATGGGA 361  
Qy 362 CCTACTGTCGAGATAGAGTTCCCTGCTCTATGATATGAATAAATAAATTAGAACTGAAT 421  
Db 362 CCTACTGTCGAGATAGAGTTCCCTGCTCTATGATATGAATAAATAAATTAGAACTGAAT 421  
Qy 422 TAGACATCAAAAGCAGCCAAAGGCTCACTCCAGCTCAAGCTGCGGAGCTCTACTACAG 481  
Db 422 TAGACATCAAAAGCAGCCAAAGGCTCACTCCAGCTCAAGCTGCGGAGCTCTACTACAG 481  
Qy 482 CTTCGCCAAGAACCTTAACACGAGAGAAAATGTTTCAGAGACAGACACACTGCTGACCC 541  
Db 482 CTTCGCCAAGAACCTTAACACGAGAGAAAATGTTTCAGAGACAGACACACTGCTGACCC 541  
Qy 542 TCATATATTAACAAATGGAACAAAATTTCCACATGGGCTGATGAATTAAGAACTCTGAG 601  
Db 542 TCATATATTAACAAATGGAACAAAATTTCCACATGGGCTGATGAATTAAGAACTCTGAG 601  
Qy 602 AAACGATAGAACTGCTATCCACATGGAAGTGGAGCTCTGCTGGGTTGACCTTGAC 661  
Db 602 AAACGATAGAACTGCTATCCACATGGAAGTGGAGCTCTGCTGGGTTGACCTTGAC 661  
Qy 662 TTATCATGTTGCTCTTAATTCCTTAATAGTATTCCTGTAAGAAAAGAAATATGAGATT 721  
Db 662 TTATCATGTTGCTCTTAATTCCTTAATAGTATTCCTGTAAGAAAAGAAATATGAGATT 721  
Qy 722 TGAGCCTTATTACATGCGCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGAGTCA 781  
Db 722 TGAGCCTTATTACATGCGCAACTTGCTCCAGAGAGGTTGGCAATGACAGAGAGTCA 781  
Qy 782 GGAATCGCTGAGAGAAAATATCTACACCATGAGAGAAAGTATATGAAAGTGAGAAATT 841  
Db 782 GGAATCGCTGAGAGAAAATATCTACACCATGAGAGAAAGTATATGAAAGTGAGAAATT 841  
Qy 842 CAAAGACTTACTACTGCTACGTCACACAGCAGACACATCCGACCGCTCTGAGCTGCC 901  
Db 842 CAAAGACTTACTACTGCTACGTCACACAGCAGACACATCCGACCGCTCTGAGCTGCC 901  
Qy 902 ACTTTAAAGGCTGCGCTTCACTTTGACCTTGATTTGCTCTTCTGGAATACTATG 961  
Db 902 ACTTTAAAGGCTGCGCTTCACTTTGACCTTGATTTGCTCTTCTGGAATACTATG 961  
Qy 962 ATATGTCATTTGCAACCTCATTTGAGGTTCTGACACAGCCACTGAGAAAAGATTCCA 1021  
Db 962 ATATGTCATTTGCAACCTCATTTGAGGTTCTGACACAGCCACTGAGAAAAGATTCCA 1021  
Qy 1022 GTTTTCGAGGATATTAATCAACAAGGGAATTGCACTGTAACTATGTAATTAAT 1081  
Db 1022 GTTTTCGAGGATATTAATCAACAAGGGAATTGCACTGTAACTATGTAATTAAT 1081  
Qy 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTCGAGATCCCACTCCAGAGACTTCAATCATG 1141  
Db 1082 GCTCCATTTTATCCCTGAGTTTCAGGAGTCGAGATCCCACTCCAGAGACTTCAATCATG 1141  
Qy 1142 CGTGTGGAAGCTCACTGCTGCTTTCAATCATTTAGGAATGTTAGTGTCTTTGAGA 1201  
Db 1142 CGTGTGGAAGCTCACTGCTGCTTTCAATCATTTAGGAATGTTAGTGTCTTTGAGA 1201  
Qy 1202 CATTAAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTATGGGGAATTAAGGCTTA 1261  
Db 1202 CATTAAGGTTTGTGTATATCCGCAAGCTCTGAAACAGTATGGGGAATTAAGGCTTA 1261  
Qy 1262 GATTAAGAGTGCCTGCTTTGTTGATGTTGAAAAATCTTAAGAAAGTTGTAGCTTTTCT 1321  
Db 1262 GATTAAGAGTGCCTGCTTTGTTGATGTTGAAAAATCTTAAGAAAGTTGTAGCTTTTCT 1321  
Qy 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAGGTGACATATGCTTAACACGATAT 1381  
Db 1322 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAGGTGACATATGCTTAACACGATAT 1381

Db 1382 AGAGATTTCTGACCTTGAAGATTAAGAAAAGCAGGTGACATATGCTTAACACGATAT 1381  
Qy 1382 AACTTGGGAACCTTAGGAGAGGTTGATTAAGTTCAAGGTGACGCGAGGCTATGCTGTA 1441  
Db 1382 AACTTGGGAACCTTAGGAGAGGTTGATTAAGTTCAAGGTGACGCGAGGCTATGCTGTA 1441  
Qy 1442 AGACTGCTCAMCATCCAAAGACGAAATAAATACATAGAGACAGAGAGGCTGAGATGA 1501  
Db 1442 AGACTGCTCAMCATCCAAAGACGAAATAAATACATAGAGACAGAGAGGCTGAGATGA 1501  
Qy 1502 GGGCTGGACAGTGAAGTGCATTTGTACAGACGAGGAATCTATATTTGATCGTAGACC 1561  
Db 1502 GGGCTGGACAGTGAAGTGCATTTGTACAGACGAGGAATCTATATTTGATCGTAGACC 1561  
Qy 1562 CCACATGAAAAGCTAGGCTGTAGAGCACTTGTATGATCTCAAGAGATGAGAGGTAA 1621  
Db 1562 CCACATGAAAAGCTAGGCTGTAGAGCACTTGTATGATCTCAAGAGATGAGAGGTAA 1621  
Qy 1622 AGGCACAACAGATCCCGGGGCTTGCGTCAAGTCAAGCTTACCTAGTGTGATCCAA 1681  
Db 1622 AGGCACAACAGATCCCGGGGCTTGCGTCAAGTCAAGCTTACCTAGTGTGATCCAA 1681  
Qy 1682 GTCCAAAGAGTCCCTGTCTCAMAAGTAAGATGACGTAGATCTGCGCATGTCATGG 1741  
Db 1682 GTCCAAAGAGTCCCTGTCTCAMAAGTAAGATGACGTAGATCTGCGCATGTCATGG 1741  
Qy 1742 GGTGTCTCTCTCTCAGAAAGACATGCAATGWCCTCTCACAACACACACACAC 1801  
Db 1742 GGTGTCTCTCTCTCAGAAAGACATGCAATGWCCTCTCACAACACACACACACAC 1801  
Qy 1802 ACAACACACACACACACACACACACACATGAAATGAAAGTTCTCTGCTGCTCTAC 1861  
Db 1802 ACAACACACACACACACACACACACACATGAAATGAAAGTTCTCTGCTGCTCTAC 1861  
Qy 1862 TCTCTAATACATGATCTCTACAGAGCTCTCTCTGCTGCTTAAAGATGAGGAG 1921  
Db 1862 TCTCTAATACATGATCTCTACAGAGCTCTCTCTGCTGCTTAAAGATGAGGAG 1921  
Qy 1922 CATGGCAGAGATCCAGTAATTTATTCACACATCAAGAGCTGAGACAGAGGTGA 1981  
Db 1922 CATGGCAGAGATCCAGTAATTTATTCACACATCAAGAGCTGAGACAGAGGTGA 1981  
Qy 1982 GAGTTCAAGAGCACTGTCGCCCAACATGCGACACTCTCTTCAACAAGAAAAGTTACC 2041  
Db 1982 GAGTTCAAGAGCACTGTCGCCCAACATGCGACACTCTCTTCAACAAGAAAAGTTACC 2041  
Qy 2042 CGCAAGCAGCTGCTGTCTGTAAGAAAGAACCTGCGAAGGCAAACTTTGACTGTGTG 2101  
Db 2042 CGCAAGCAGCTGCTGTCTGTAAGAAAGAACCTGCGAAGGCAAACTTTGACTGTGTG 2101  
Qy 2102 TGCTCAAGGGAATGATCTGACACAACTTCTCATTTCTGAGAGAAACTGAGAGCTGTTT 2161  
Db 2102 TGCTCAAGGGAATGATCTGACACAACTTCTCATTTCTGAGAGAAACTGAGAGCTGTTT 2161  
Qy 2162 TGACAGAAAGCAACCGGTGACTGGGACATACGAAGCAGAGCTTTGACAGAACTCTTA 2221  
Db 2162 TGACAGAAAGCAACCGGTGACTGGGACATACGAAGCAGAGCTTTGACAGAACTCTTA 2221  
Qy 2222 TAGTCAGCAAAATATTTCTTGGAGAGACGTGTCACCAATTTGATTTCAAGCCGGTGG 2281  
Db 2222 TAGTCAGCAAAATATTTCTTGGAGAGACGTGTCACCAATTTGATTTCAAGCCGGTGG 2281  
Qy 2282 AACTCAGTTTCAATGCTGCTTACAGTGTGCTGCGCCAGTGTATCTGTGCTGCTGCC 2341  
Db 2282 AACTCAGTTTCAATGCTGCTTACAGTGTGCTGCGCCAGTGTATCTGTGCTGCTGCC 2341  
Qy 2342 ATCTATAACAGAAATCAAAATTAATAGACCCGAGTGAATAATTAAGTGAAGCAAGAT 2401  
Db 2342 ATCTATAACAGAAATCAAAATTAATAGACCCGAGTGAATAATTAAGTGAAGCAAGAT 2401  
Qy 2402 AGCTTTGTTCAAGATTTTCTTGAATGGAGACACTGTGTACATCAGAGACACTGTGT 2461  
Db 2402 AGCTTTGTTCAAGATTTTCTTGAATGGAGACACTGTGTACATCAGAGACACTGTGT 2461

QY 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTCATGTATGAAATTTGTGTAGGT 2521  
DB 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTCATGTATGAAATTTGTGTAGGT 2521  
QY 2522 TGCTTACTAGCTGTGTAGAGTCTGTGGCTTCTTGTAGGTGTATGAAAGGAGACCATC 2581  
DB 2522 TGCTTACTAGCTGTGTAGAGTCTGTGGCTTCTTGTAGGTGTATGAAAGGAGACCATC 2581  
QY 2582 TAACAAAATCCATTAGATTAACAGCTCTCATGACAGAGGAAACTAAATCTCAATGTT 2641  
DB 2582 TAACAAAATCCATTAGATTAACAGCTCTCATGACAGAGGAAACTAAATCTCAATGTT 2641  
QY 2642 TTTAAATTAATAAAGTGTGTGTGCAAAAGTCTTGTGAGCAATTAATAAATTAATAAAG 2701  
DB 2642 TTTAAATTAATAAAGTGTGTGTGCAAAAGTCTTGTGAGCAATTAATAAATTAATAAAG 2701  
QY 2702 GGGGGCCGC 2710  
DB 2702 GGGGGCCGC 2710

RESULT 5  
US-08-937-399-8  
Sequence 8, Application US/08937399  
Patent No. 6288218  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,399  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 40..885  
US-08-937-399-8

Query Match 99.9%; Score 2706.6; DB 3; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGACCCCAAGGCGTCGGATTTCCCTCCCAAGTACTATGTTTTCAGGTCTTAACCTCA 61  
DB 2 GTGACCCCAAGGCGTCGGATTTCCCTCCCAAGTACTATGTTTTCAGGTCTTAACCTCA 61  
QY 62 ACTGTGCTGTGCTGTGCAACTACTACTTGAAGGTCAATGGAATGTTATAGG 121  
DB 62 ACTGTGCTGTGCTGTGCAACTACTACTTGAAGGTCAATGGAATGTTATAGG 121  
QY 122 TTGAGTTGTATTAATAATGCTATCTGCTGTGAGTTACCTTACTTATGGAACAC 181  
DB 122 TTGAGTTGTATTAATAATGCTATCTGCTGTGAGTTACCTTACTTATGGAACAC 181  
QY 182 TTGTGCTATGTGTGTGGGCAAGGATTTCTGCTTGTGTCAGAGTATCAATGAGTTGC 241  
DB 182 TTGTGCTATGTGTGTGGGCAAGGATTTCTGCTTGTGTCAGAGTATCAATGAGTTGC 241  
QY 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGAGATACAGTTAAAGGCG 301  
DB 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGAGATACAGTTAAAGGCG 301  
QY 302 ATCTCAACAAAGAGATGTCTGTGATCATTAAGAAATGTGATGATGATGATGATGATG 361  
DB 302 ATCTCAACAAAGAGATGTCTGTGATCATTAAGAAATGTGATGATGATGATGATGATG 361  
QY 362 CCTACTGTGACAGATACAGTTCCCTGTGCTTATGATGATGATGATGATGATGATGATG 421  
DB 362 CCTACTGTGACAGATACAGTTCCCTGTGCTTATGATGATGATGATGATGATGATGATG 421  
QY 422 TAGACATCAAGACGCAAGGTCTGCTGAGTCAAGTCCAGTCCAGTCCAGTCCAGTCCAG 481  
DB 422 TAGACATCAAGACGCAAGGTCTGCTGAGTCAAGTCCAGTCCAGTCCAGTCCAGTCCAG 481  
QY 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTGATGATGATGATGATGATGATGATG 541  
DB 482 CTCTCCCAAGAACCTTAACCAAGGAGAGAAATGTGATGATGATGATGATGATGATGATG 541  
QY 542 TCCATATATACAGTGAACAAATTTCCACATGAGGCTGATGATGATGATGATGATGATG 601  
DB 542 TCCATATATACAGTGAACAAATTTCCACATGAGGCTGATGATGATGATGATGATGATG 601  
QY 602 AAACGATCAGAACTGCTATCCACATGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTG 661  
DB 602 AAACGATCAGAACTGCTATCCACATGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTG 661  
QY 662 TTAATCATGCTGCTTAATCCCTTAATGATTTCCCTGTAAGAAAGAAAGTATGAGTT 721  
DB 662 TTAATCATGCTGCTTAATCCCTTAATGATTTCCCTGTAAGAAAGAAAGTATGAGTT 721  
QY 722 TGAGCTTATTAACAGTGGCACTTGCTCCAGAGAGGTTGGCAATGAGAGAGAGTCA 781  
DB 722 TGAGCTTATTAACAGTGGCACTTGCTCCAGAGAGGTTGGCAATGAGAGAGAGTCA 781  
QY 782 GGAATGCTCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGGAGAA 841  
DB 782 GGAATGCTCTGAGAGAAATATCTACACCATGAGAGAAAGTATATGAGTGGAGAA 841  
QY 842 CAAATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
DB 842 CAAATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
QY 902 ACTTTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961  
DB 902 ACTTTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961  
QY 962 ATATGCTACTGAGCAACCTCATTTGAGAGGTTGAGACCAAGCACTGAGAAAGAGTTCCA 1021  
DB 962 ATATGCTACTGAGCAACCTCATTTGAGAGGTTGAGACCAAGCACTGAGAAAGAGTTCCA 1021











Dh 302 ATCTCAACAAGAGATGTCTCTGATCAATAAGATGTGACTCTGATGACCAATGGCA 361  
Qy 362 CCTACTGCTGCAGAGATACAGTTCCCTGGTCTTAATGATGATTAATAAATTAGAACTGAAT 421  
Dh 362 CCTACTGCTGCAGAGATACAGTTCCCTGGTCTTAATGATGATTAATAAATTAGAACTGAAT 421  
Qy 422 TAGACATCAAAAGACCCAAAGSTCACTCCAGCTCAGACTGCGCCATGGGAGCTCTACTACAG 481  
Dh 422 TAGACATCAAAAGACCCAAAGSTCACTCCAGCTCAGACTGCGCCATGGGAGCTCTACTACAG 481  
Qy 482 CTCTCTCAAGAACCTTACCAAGGAGAAATGATGATCAGAGACACAGACTGATGAGCC 541  
Dh 482 CTCTCTCAAGAACCTTACCAAGGAGAAATGATGATCAGAGACACAGACTGATGAGCC 541  
Qy 542 TCCATATAATCAATGGAACAAAATTTCCATGAGGCTGATGAAATTTAGAGACTCTGAG 601  
Dh 542 TCCATATAATCAATGGAACAAAATTTCCATGAGGCTGATGAAATTTAGAGACTCTGAG 601  
Qy 602 AAACGATCAGAACTGCTATCCACTTGGAGTGGAGTCTCTGCTGGTTGACCTGGAC 661  
Dh 602 AAACGATCAGAACTGCTATCCACTTGGAGTGGAGTCTCTGCTGGTTGACCTGGAC 661  
Qy 662 TTATCATGCTGTCTTATCTTAAATGATGATCTGTAGAGAAAGAGATTAATGAGTT 721  
Dh 662 TTATCATGCTGTCTTATCTTAAATGATGATCTGTAGAGAAAGAGATTAATGAGTT 721  
Qy 722 TGAGCTTATTAACATGAGCACTGGCACTTGGCTCAGAGAGGTTGGCAATGAGAGCACTCA 781  
Dh 722 TGAGCTTATTAACATGAGCACTGGCACTTGGCTCAGAGAGGTTGGCAATGAGAGCACTCA 781  
Qy 782 GGAATTCGCTCTGAGGAAATATCTACACCATCCAGAGGAAACGTAATAGATGAGAAAT 841  
Dh 782 GGAATTCGCTCTGAGGAAATATCTACACCATCCAGAGGAAACGTAATAGATGAGAAAT 841  
Qy 842 CAATGAGTACTACTGCTGATGATCAACAGCCAGCAGCATCTGACCGCTCTGAGCTGCC 901  
Dh 842 CAATGAGTACTACTGCTGATGATCAACAGCCAGCAGCATCTGACCGCTCTGAGCTGCC 901  
Qy 902 ACTTTTAAAGGCTCGCTTCAATTTCTGACTTTGGTATTTCCCTTTTGGAAACATATG 961  
Dh 902 ACTTTTAAAGGCTCGCTTCAATTTCTGACTTTGGTATTTCCCTTTTGGAAACATATG 961  
Qy 962 ATATGTCATCTGGCACTGATGAGGATCTGACCAAGCCACTGAGCACTGAGAAAGATTC 1021  
Dh 962 ATATGTCATCTGGCACTGATGAGGATCTGACCAAGCCACTGAGCACTGAGAAAGATTC 1021  
Qy 1022 GTTTTCTGGGATATTAATCACTCAAGGGGATTCGATGATCACTGATCACTGATCACT 1081  
Dh 1022 GTTTTCTGGGATATTAATCACTCAAGGGGATTCGATGATCACTGATCACTGATCACT 1081  
Qy 1082 GCTTCATTTATCTCGAGTTTCAAGGATCGGATCTCCCACTCAGAGACTTCATCATG 1141  
Dh 1082 GCTTCATTTATCTCGAGTTTCAAGGATCGGATCTCCCACTCAGAGACTTCATCATG 1141  
Qy 1142 CGTGTGAAGCTCACTGCTCTTCACTCAATAGAAATGATGATGATGATGATGATGATG 1201  
Dh 1142 CGTGTGAAGCTCACTGCTCTTCACTCAATAGAAATGATGATGATGATGATGATGATG 1201  
Qy 1202 CATAGAGTTTGTGATATCCGCAAGCTCTGAAACAGGTAAGGGGATTAAGGGGATTA 1261  
Dh 1202 CATAGAGTTTGTGATATCCGCAAGCTCTGAAACAGGTAAGGGGATTAAGGGGATTA 1261  
Qy 1262 GATAGAGAGTGGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1321  
Dh 1262 GATAGAGAGTGGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1321  
Qy 1322 AGAGATTTCTGACTTAAAGATTAAGAAAAAGCAGAGTGGCATATCTTAACAGAT 1381  
Dh 1322 AGAGATTTCTGACTTAAAGATTAAGAAAAAGCAGAGTGGCATATCTTAACAGAT 1381  
Qy 1382 AACTTGGGAACCTTAGGCAAGAGGATGATTAAGTTCAAGGTCAGGCAAGGCTATGCTGTA 1441  
Dh 1382 AACTTGGGAACCTTAGGCAAGAGGATGATTAAGTTCAAGGTCAGGCAAGGCTATGCTGTA 1441

Qy 1442 AGACTGTCTCAGCATCCAAAGAGAAATTAACATAGAGACAGCAGAGGCTGGAGATGA 1501  
Dh 1442 AGACTGTCTCAGCATCCAAAGAGAAATTAACATAGAGACAGCAGAGGCTGGAGATGA 1501  
Qy 1502 GGTGGGACAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1561  
Dh 1502 GGTGGGACAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1561  
Qy 1562 CCACATGAAAAAGCTGAGGCTGATGAGATGATGATGATGATGATGATGATGATGATGATG 1621  
Dh 1562 CCACATGAAAAAGCTGAGGCTGATGAGATGATGATGATGATGATGATGATGATGATGATG 1621  
Qy 1622 AGGACACAGATCCCGGGGCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 1681  
Dh 1622 AGGACACAGATCCCGGGGCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 1681  
Qy 1682 GTCCACAAAGTCCCTGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1741  
Dh 1682 GTCCACAAAGTCCCTGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1741  
Qy 1742 GGTGTCTCTCTCTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
Dh 1742 GGTGTCTCTCTCTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
Qy 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
Dh 1802 ACAACACACACACACACACACACACACATGATGATGATGATGATGATGATGATGATGATG 1861  
Qy 1862 TCTCTTAAATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1921  
Dh 1862 TCTCTTAAATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1921  
Qy 1922 CATGAGAGAGAGTCCAGATATTTATTCAGACCTCAGAGAGCTGAGAGAGAGCTGGA 1981  
Dh 1922 CATGAGAGAGAGTCCAGATATTTATTCAGACCTCAGAGAGCTGAGAGAGAGCTGGA 1981  
Qy 1982 GAGTTCAAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2041  
Dh 1982 GAGTTCAAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2041  
Qy 2042 CGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
Dh 2042 CGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101  
Qy 2102 TGCTCAAGGGGAACTGATCAAGCACTTCTCATCTCTGAGAGAACTGAGCTGTTTC 2161  
Dh 2102 TGCTCAAGGGGAACTGATCAAGCACTTCTCATCTCTGAGAGAACTGAGCTGTTTC 2161  
Qy 2162 TGACAGAAAGCAACCGGTGATGAGGACATAGAGGAGAGCTTGGAGCAATCTTA 222  
Dh 2162 TGACAGAAAGCAACCGGTGATGAGGACATAGAGGAGAGCTTGGAGCAATCTTA 222  
Qy 2222 TAGTCAGCAAAATATCTTGGGAGAGAGTGCATCAAAATGATTTCAAGCCGAGTGG 2281  
Dh 2222 TAGTCAGCAAAATATCTTGGGAGAGAGTGCATCAAAATGATTTCAAGCCGAGTGG 2281  
Qy 2282 ACCTCAGTTTCACTGCTTCAAGCTGCTGCCAGTGCCTTGAATCTGCTGCTGCC 2341  
Dh 2282 ACCTCAGTTTCACTGCTTCAAGCTGCTGCCAGTGCCTTGAATCTGCTGCTGCC 2341  
Qy 2342 ATCTATTAACAGATCAAAATTAATAGACCCGAGTAAAAATTTAGTGAAGCAAGG 2401  
Dh 2342 ATCTATTAACAGATCAAAATTAATAGACCCGAGTAAAAATTTAGTGAAGCAAGG 2401  
Qy 2402 AGCTTTGTTCAAAAGATTTTTCATGATGAGGAGCACTGATCAAGAGCATCTG 2461  
Dh 2402 AGCTTTGTTCAAAAGATTTTTCATGATGAGGAGCACTGATCAAGAGCATCTG 2461  
Qy 2462 TAGTGAAGACACCAAAACCTGATGATCCGTTTTCATGATGATGATGATGATGATGATG 2521  
Dh 2462 TAGTGAAGACACCAAAACCTGATGATCCGTTTTCATGATGATGATGATGATGATGATG 2521

QY 2522 TCGCTTACGACGCTGCTGAGGCTCTGCTTCTTAGTGCGTATGAGAGGAGACATC 2581  
DB 2522 TCGCTTACGACGCTGCTGAGGCTCTGCTTCTTAGTGCGTATGAGAGGAGACATC 2581  
QY 2582 TACCAAAATCCATTAGAGTAAACAGCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
DB 2582 TACCAAAATCCATTAGAGTAAACAGCTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTAAGTATATAAAGCTGCTGCGCAAGTACTTTGAGCATATAAAAAAAAAAAG 2701  
DB 2642 TTAAGTATATAAAGCTGCTGCGCAAGTACTTTGAGCATATAAAAAAAAAAAG 2701  
QY 2702 GCGGCGCGC 2710  
DB 2702 GCGGCGCGC 2710

RESULT 8  
US-09-464-231-8  
; Sequence 8, Application US/09464231  
; Patent No. 6562343  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/464,231  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cornuzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-048  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNTE  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2710 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 40..885  
US-09-464-231-8

Query Match 99.9%; Score 2706.6; DB 4; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGACACCGCGTCCGAGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCA 61  
DB 2 GTGACACCGCGTCCGAGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTACCTCA 61  
QY 62 ACTGTGCTCTGCTGCTGCACTACTACTTTCAGAGGTCAATGGATTTAAAG 121  
DB 62 ACTGTGCTCTGCTGCTGCACTACTACTTTCAGAGGTCAATGGATTTAAAG 121  
QY 122 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACAGTAACTACTATCTGAGAAC 181  
DB 122 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACAGTAACTACTATCTGAGAAC 181  
QY 182 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACAGTAACTACTATCTGAGAAC 241  
DB 182 TTGAGGTGGTAAAAATCCCTATCTGCCCCGACAGTAACTACTATCTGAGAAC 241  
QY 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGACAGATACAGCTAAAGGCG 301  
DB 242 TCAGAACTGATGAAAGAAATGTGACATATCAGAAATCCAGACAGATACAGCTAAAGGCG 301  
QY 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTGACTCTGATACCATGGGA 361  
DB 302 ATCTCAACAAAGAGATGTCTCTGATCATTAAGAAATGTGACTCTGATACCATGGGA 361  
QY 362 CCTACTGCTGACAGATACAGTTCCTGCTTATGATGATGATGATGATGATGATGATGAT 421  
DB 362 CCTACTGCTGACAGATACAGTTCCTGCTTATGATGATGATGATGATGATGATGATGAT 421  
QY 422 TAGACATCAAGACAGCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 481  
DB 422 TAGACATCAAGACAGCAGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 481  
QY 482 CTCTCCCAAGAACCTTACACAGAGAGAAATGTGACAGACACAGACACAGCTGTGACCC 541  
DB 482 CTCTCCCAAGAACCTTACACAGAGAGAAATGTGACAGACACAGACACAGCTGTGACCC 541  
QY 542 TCCATATATCAATGAGCAAAATTTTCCATAGGCTGATGATGATGATGATGATGATGATG 601  
DB 542 TCCATATATCAATGAGCAAAATTTTCCATAGGCTGATGATGATGATGATGATGATGATG 601  
QY 602 AAACGATCAGAACTGCTATCCACATTTGAGAGGAGTCTGCTGCTGCTGCTGCTGCTG 661  
DB 602 AAACGATCAGAACTGCTATCCACATTTGAGAGGAGTCTGCTGCTGCTGCTGCTGCTG 661  
QY 662 TTATCATTTGCTTAAATCTTAAATGATGATGATGATGATGATGATGATGATGATG 721  
DB 662 TTATCATTTGCTTAAATCTTAAATGATGATGATGATGATGATGATGATGATGATG 721  
QY 722 TGAGCCTTATTAACATGAGCACTGCTCCAGAGAGGTTGGCAAAATGACAGAGCAGTCA 781  
DB 722 TGAGCCTTATTAACATGAGCACTGCTCCAGAGAGGTTGGCAAAATGACAGAGCAGTCA 781  
QY 782 GGATTCGCTGAGGAAATATCTACACATGAGAGAGAGATGATGATGATGATGATGATG 841  
DB 782 GGATTCGCTGAGGAAATATCTACACATGAGAGAGAGATGATGATGATGATGATGATG 841  
QY 842 CAAATGATTACTACTGCTACGTCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 901  
DB 842 CAAATGATTACTACTGCTACGTCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 901  
QY 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTGTGATTTCTGCTTGTGAAATCTATG 961  
DB 902 ACTTTTAAAGGCTGCTTCAATTTCTGACTTGTGATTTCTGCTTGTGAAATCTATG 961  
QY 962 ATATGCTACTTGGCACTCAATTTGAGAGGTTTGAACACAGCAGCAGCAGCAGCAGCAG 1021  
DB 962 ATATGCTACTTGGCACTCAATTTGAGAGGTTTGAACACAGCAGCAGCAGCAGCAGCAG 1021  
QY 1022 GTTTTCGAGATATATTAATCAACAAGGAGTGTGACGTATGATGATGATGATGATGAT 1081  
DB 1022 GTTTTCGAGATATATTAATCAACAAGGAGTGTGACGTATGATGATGATGATGATGAT 1081  
QY 1082 GCTCATTTTATCCCTGAGTTTCAAGGATGAGATGATGATGATGATGATGATGATGATG 1141

DB 1082 GCTCCATTATATCCGAGTTTCAGGAGTGGATCTCCACTCCAGGACTTCAATCATG 1141  
QY 1142 CGTGTGAAGCTACTGTCCTTTCATACATTAAGATAGTGTATGATGTCCTTGAGA 1201  
DB 1142 CGTGTGAAGCTACTGTCCTTTCATACATTAAGATAGTGTATGATGTCCTTGAGA 1201  
QY 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAAACAGTAAAGGAGGCTAA 1261  
DB 1202 CATAGAGTTTGTGTATATCCGAAAGCTCTGAAACAGTAAAGGAGGCTAA 1261  
QY 1262 GATAGGAAGGTGGGCTTTGTGTATGATGTTGAAATCTTAAAGAGTTGTAGCTTTCT 1321  
DB 1262 GATAGGAAGGTGGGCTTTGTGTATGATGTTGAAATCTTAAAGAGTTGTAGCTTTCT 1321  
QY 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAAGCAGGTGCAATGCTTTAACAAGATAT 1381  
DB 1322 AGAGATTTCTGACCTTGAAAGATTAAGAAAAAGCAGGTGCAATGCTTTAACAAGATAT 1381  
QY 1382 AACTTGGGAACTTTAGGACAGAGGTGATTAAGTTCAAGGTCAGGCGCTATGCTGTA 1441  
DB 1382 AACTTGGGAACTTTAGGACAGAGGTGATTAAGTTCAAGGTCAGGCGCTATGCTGTA 1441  
QY 1442 AGACTGTCTGACCTTGAAAGATTAAGAAAAAGCAGGTGCAATGCTTTAACAAGATAT 1501  
DB 1442 AGACTGTCTGACCTTGAAAGATTAAGAAAAAGCAGGTGCAATGCTTTAACAAGATAT 1501  
QY 1502 GGGTCGGAACAGTGAAGTGTATGTAACAAGCAAGAGAACTATATTTATGCTAGACC 1561  
DB 1502 GGGTCGGAACAGTGAAGTGTATGTAACAAGCAAGAGAACTATATTTATGCTAGACC 1561  
QY 1562 CCACATGAAAAAGCTAGGCTGTGTAAGCAATGCTTGAAGCTCAAGAGTGAAGGTAA 1621  
DB 1562 CCACATGAAAAAGCTAGGCTGTGTAAGCAATGCTTGAAGCTCAAGAGTGAAGGTAA 1621  
QY 1622 AGGCAACAAGATCCCGGGGGCTTGGCTGAGTCAAGCTTGAAGGTCGAGTTCCAA 1681  
DB 1622 AGGCAACAAGATCCCGGGGGCTTGGCTGAGTCAAGCTTGAAGGTCGAGTTCCAA 1681  
QY 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGTGAAGTATCTGGGCAATGTCATAGG 1741  
DB 1682 GTCCACAAGAGTCCCTGTCTCAMAAGTAAGTGAAGTATCTGGGCAATGTCATAGG 1741  
QY 1742 GGTGTCTCTCTCTCTCTCAAGAGACATGCAATGACCTTGCACACACACACACAC 1801  
DB 1742 GGTGTCTCTCTCTCTCTCAAGAGACATGCAATGACCTTGCACACACACACACAC 1801  
QY 1802 ACACACACACACACACACACACACACATGAAATGAAAGTTCTCTGTGCTGCTAAC 1861  
DB 1802 ACACACACACACACACACACACACACATGAAATGAAAGTTCTCTGTGCTGCTAAC 1861  
QY 1862 TCTCTATATACATGATATCTTACAGAGTCTCTCTGCTCTGTTAAGACATGAGTGGAG 1921  
DB 1862 TCTCTATATACATGATATCTTACAGAGTCTCTCTGCTCTGTTAAGACATGAGTGGAG 1921  
QY 1922 CATGGCAGACATGCTCAATTTATTTTCCAGCACTCAAGAGGCTGAGCAGAAAGCTGGA 1981  
DB 1922 CATGGCAGACATGCTCAATTTATTTTCCAGCACTCAAGAGGCTGAGCAGAAAGCTGGA 1981  
QY 1982 GAGTTCAAGAGCACTGTGCCAACATGCGACATCTTCTTACACAGAAAAAGTTAAC 2041  
DB 1982 GAGTTCAAGAGCACTGTGCCAACATGCGACATCTTCTTACACAGAAAAAGTTAAC 2041  
QY 2042 CGCAAGAGCTGTGCTGTCTGTAAGAGAAACCTTGCAAGAGCAAACTTGAATGTTGTG 2101  
DB 2042 CGCAAGAGCTGTGCTGTCTGTAAGAGAAACCTTGCAAGAGCAAACTTGAATGTTGTG 2101  
QY 2102 TGCTCAAGGGGAACTGACACAACTTCTTCTTCTGAGAGAACTGAGCTGTTTC 2161  
DB 2102 TGCTCAAGGGGAACTGACACAACTTCTTCTTCTGAGAGAACTGAGCTGTTTC 2161  
QY 2162 TGACAGAAAGCAACCGGTGATGCTGGGACATACGAAGGAGAGCTTTCGACGACATCTATA 2221

DB 2162 TGACAGAAAGCAACCGGTGACTGGGACATACGAAGCAGAGCTTTCGACGACATCTATA 2221  
QY 2222 TAGTCAGAAAATATTTCTTTGGAGAGACAGTCTGCACAAATTTGATTTTCAAGCCGTGG 2281  
DB 2222 TAGTCAGAAAATATTTCTTTGGAGAGACAGTCTGCACAAATTTGATTTTCAAGCCGTGG 2281  
QY 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGTGCCTTGATCTGTGCTGCC 2341  
DB 2282 ACCTCAGTTTCATCTGGCTTACAGCTGCTGCCAGTGCCTTGATCTGTGCTGCC 2341  
QY 2342 ATCTATACAGATCAATTTAATAGACCCGAGTGAATATTAAGTAGAGCAAGAAAGT 2401  
DB 2342 ATCTATACAGATCAATTTAATAGACCCGAGTGAATATTAAGTAGAGCAAGAAAGT 2401  
QY 2402 AGCTTTGTTCAAGATTTTTCATGATGGGAGCAACTGTGTACATCAAGACATCTGT 2461  
DB 2402 AGCTTTGTTCAAGATTTTTCATGATGGGAGCAACTGTGTACATCAAGACATCTGT 2461  
QY 2462 TAGTGAGACACCAACCTGTGTACCGTTTTCATGATGAATTTGTTTAAAGT 2521  
DB 2462 TAGTGAGACACCAACCTGTGTACCGTTTTCATGATGAATTTGTTTAAAGT 2521  
QY 2522 TGCTTCTAGCTAGCTGTGAGGCTCTGCTTTCTTATGATGGTATGAAAGGAGACATC 2581  
DB 2522 TGCTTCTAGCTAGCTGTGAGGCTCTGCTTTCTTATGATGGTATGAAAGGAGACATC 2581  
QY 2582 TAACAAAATCATTAGAGATTAACAGCTTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
DB 2582 TAACAAAATCATTAGAGATTAACAGCTTTCATGACAGAGGAAACTAATCTCAATGTT 2641  
QY 2642 TTTAAGTAAATTAACGTGACTGGCAAGTACTTTGACATTAAGAAAAAAG 2701  
DB 2642 TTTAAGTAAATTAACGTGACTGGCAAGTACTTTGACATTAAGAAAAAAG 2701  
QY 2702 GGGGGCCGC 2710  
DB 2702 GGGGGCCGC 2710

RESULT 9  
US-08-829-525-22  
; Sequence 22, Application US/08829525  
; Patent No. 6084083  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITILE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Penite & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/829,525  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/609,583  
; FILING DATE: 01-MAR-1996  
; APPLICATION NUMBER: US 08/487,748  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/398,633  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.











```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,337
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,583
FILING DATE: 01-MAR-1996
APPLICATION NUMBER: US 08/487,748
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/398,633
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-032-337-22

Query Match      31.1%; Score 843; DB 4; Length 843;
Best Local Similarity 100.0%; Pred. No. 3e-235;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 580 GATGAATTAAGAGCTGTGGAGAAAGATCAGACTGCTATCCATGAGTGGAGATC 639
DB 541 GATGAATTAAGAGCTGTGGAGAAAGATCAGACTGCTATCCATGAGTGGAGATC 600
QY 640 TCTGCTGGGTGACCTGTGACCTTATCATGAGTGTCTTATATGATGATTCCTGT 699
DB 601 TCTGCTGGGTGACCTGTGACCTTATCATGAGTGTCTTATATGATGATTCCTGT 660
QY 700 AAGAAAAGAAAGTATTCAGTTGAGCTTATTAACCTGCGCACTGCTCCAGAGG 759
DB 661 AAGAAAAGAAAGTATTCAGTTGAGCTTATTAACCTGCGCACTGCTCCAGAGG 720
QY 760 TTGGCAAAATGACAGAGCACTGAGATGCTGCTGAGAAAATATCTACACATCGAGAG 819
DB 721 TTGGCAAAATGACAGAGCACTGAGATGCTGCTGAGAAAATATCTACACATCGAGAG 780
QY 820 AACGTATATGAGTGGAGATTCATTAATGATGATCTACTGCTACGTCACAGCAGACCA 879
DB 781 AACGTATATGAGTGGAGATTCATTAATGATGATCTACTGCTACGTCACAGCAGACCA 840
QY 880 TCC 882
DB 841 TCC 843
```

```
RESULT 14
US-09-464-231-22
Sequence 22, Application US/09464231
Patent No. 6562343
GENERAL INFORMATION:
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESSES:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/464,231
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/609,583
FILING DATE: 01-MAR-1996
APPLICATION NUMBER: US 08/487,748
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/398,633
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-464-231-22
```



Db 316 ATGTGTCCCTGACCATAGAGATGTGACTCTAGACAGACAGTGGGATCTACTGCTGCCGA 375  
 QY 377 TACAGTTCCTGTGCTTATGATGATGATAAAATTAGAACTGAAATTAGACATCAAAAGCAG 436  
 Db 376 TCCAAATCCAGGACATATGATGATAAAATTAACTGAAGTTGTCATCAAAACAG 435  
 QY 437 CCAAGGTCACTCCAGCTCAGACTGCCCATGGGAGCTCTACTACAGTTCCTCAAGAACCC 496  
 Db 436 CCAAGGTCACTCCAGCTCAGACTGCCCATGGGAGCTCTACTACAGTTCCTCAAGAGTGC 495  
 QY 497 TAAACAAGAGAGAAATG---TTCAAGACACAGACACTGTGACCCCTCATATATACA 553  
 Db 496 TTACCAACAGGAGACATGGCCACAGAGACACAGACACTGGGAGGCTCCCTGATATA 555  
 QY 554 ATGGAACAAAATTTCAC-----ATGGCTGATGAAA 586  
 Db 556 ATCTAACACAAATATCCACATTTGGCCATGATTTACGGGACTCTAGATTGGCCAAATGACT 615  
 QY 587 TTAAAGACTCTGAGAAACGATGAGAACTGCTATCCACTTTGGAGTGGAGTCTGCTG 646  
 Db 616 TACGGACTCTGAGCAACATGAGAAATAGCATCTACATCGAGACAGGAGTCTGCTG 675  
 QY 647 GGTGACCTGAGCACTTATCATTTGCTTAAATCCTTAATGATTCCTGTAAGAAA 706  
 Db 676 GGCTGGCTGCTCTTATCTTGGGCTTTAATTTCAATGATTCCTCATAGCAAG 735  
 QY 707 AGAAGTTATGAGTTGAGCTTATTAACCTGCGCAACTTGCTCCAGAGGGTTGGCAA 766  
 Db 736 AGAAGATACAGAAATTAAGCTCATCTTTGGCCAACTCCCTCCTCAGAAATGGCAA 795  
 QY 767 ATGCAGAGAGTCAGAGATTGCTGAGAGAAATATCTACACATCGAGAGAGAGCTAT 826  
 Db 796 ATGCAGAGAGAGAGAGATTGCTCAGAGAAACATCTATACATTAAGAGAGAGCTAT 855  
 QY 827 ATGAAGTGAAGATTCAATGAGTACTGCTACGTCACAGCCAGCCACATCTCTGAC 886  
 Db 856 ATGAAGTGAAGAGCCCAATGAGTATATTGCTATGTCAGCAGAGCAGCAACCTCTAC 915  
 QY 887 CGCTCTGAGCTGCCACTTTTAAAGC-----TCGCTTCAATTTCTGACTTTGG 935  
 Db 916 AACCTTGGGTTGCTGCTTGGCAATGCAATGATCCAAACCACTTATTTTGAAGCTTGGT 975  
 QY 936 TATTCCTTTKAGAAAACTATGATGATATGTCATTTGGCAACCTCATTTGAGTTCTGA 995  
 Db 976 GTTTTGTCTTTTCAAGAACTATGAGCTGTGTCACTGCTGTT--TTGAGGTTCTGT 1033  
 QY 996 CCAACAGCACTGAGAAAGAGTTCCAGTTTCTGGGATATTAATTAATCAAGGGGATTC 1055  
 Db 1034 CCACTGTCTATGAGCAGAGTTTCCCATTTTCAGAAATAA--TGACTCACATGGGAATTG 1092  
 QY 1056 GACTGTACTCATGCTACATGAAATGCTCCATTTATCCCTGAGTTTCAAGG--ATCG 1112  
 Db 1093 AACCTGGGACTGCACTGAATTAACAGGATGTCATTTGCTGTATTTAAGCCACAG 1152  
 QY 1113 GATCTCCCACTCCAGAGACT--TCAATCATGCTGTGAAGTCACTGCTGCTT--CATAT 1169  
 Db 1153 AGTTACCAACCAAGACATGTTAATCATGAGTTAGAGCTCAAAACGGCTTTTATATA 1212  
 QY 1170 CATTAAGATGTTAGTGTATGCTTTG 1198  
 Db 1213 CACTAGGAATTTCTTGAAGTGGGCTCTG 1241

Search completed: November 22, 2003, 01:50:20  
 Job time : 176.515 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 23:34:40 ; Search time 871.58 Seconds  
(without alignments)  
10161.809 Million cell updates/sec

Title: US-10-004-633-8

Perfect score: 2710

Sequence: 1 ngtcgaccacgcgcgcga.....aaaaaaaaagcgccgcgc 2710

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

Result No.	Score	Query Match	Length	DB ID	Description
1	2706.6	99.9	2710	US-10-252-131-8	Sequence 8, Appli
2	2706.6	99.9	2710	US-10-004-633-8	Sequence 8, Appli
3	2624.6	96.8	2725	US-10-188-012-10	Sequence 10, Appli
4	844.4	31.2	862	US-10-188-012-12	Sequence 12, Appli
5	843	31.1	843	US-10-252-131-32	Sequence 22, Appli
6	843	31.1	843	US-10-004-633-22	Sequence 22, Appli
7	499.8	18.4	2236	US-10-252-131-23	Sequence 0, Appli
8	499.8	18.4	2236	US-10-004-633-23	Sequence 23, Appli
9	485.8	17.9	1116	US-10-188-012-30	Sequence 30, Appli
10	485.8	17.9	1116	US-10-188-012-32	Sequence 32, Appli
11	456	16.8	903	US-10-252-131-37	Sequence 37, Appli
12	456	16.8	903	US-10-004-633-37	Sequence 37, Appli
13	328.8	12.1	2091	US-09-529-063-36	Sequence 36, Appli
14	328.8	12.1	2091	US-10-414-378-36	Sequence 36, Appli
15	289	10.7	555	US-09-529-063-35	Sequence 35, Appli
16	289	10.7	555	US-10-414-378-35	Sequence 35, Appli

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

17	263.4	9.7	529	13	US-10-040-739-415	Sequence 415, App
18	231.6	8.5	1203	10	US-09-966-546-21	Sequence 21, Appli
19	231.6	8.5	1203	10	US-09-966-545-21	Sequence 21, Appli
20	231.6	8.5	1203	11	US-09-965-212-21	Sequence 21, Appli
21	231.6	8.5	1203	15	US-10-189-940-21	Sequence 21, Appli
22	182.6	6.7	375	15	US-10-189-940-144	Sequence 144, App
23	84.2	3.1	1079	14	US-10-188-012-24	Sequence 24, Appli
24	84.2	3.1	1080	14	US-10-188-012-18	Sequence 18, Appli
25	84.2	3.1	1080	14	US-10-188-012-20	Sequence 20, Appli
26	84.2	3.1	1095	14	US-10-188-012-26	Sequence 26, Appli
27	84.2	3.1	1099	14	US-10-188-012-22	Sequence 22, Appli
28	84.2	3.1	1099	14	US-10-188-012-28	Sequence 28, Appli
29	78.2	2.9	1032	14	US-10-188-012-14	Sequence 14, Appli
30	78.2	2.9	1032	14	US-10-188-012-16	Sequence 16, Appli
31	77.8	2.9	918	14	US-10-188-012-2	Sequence 2, Appli
32	76.8	2.8	849	14	US-10-188-012-4	Sequence 4, Appli
33	76.8	2.8	958	14	US-10-188-012-6	Sequence 6, Appli
34	76.8	2.8	958	14	US-10-188-012-8	Sequence 8, Appli
35	70.8	2.6	398	10	US-09-960-352-7732	Sequence 7732, App
36	70	2.6	418	10	US-09-960-352-1063	Sequence 1063, App
37	65.6	2.4	571	9	US-09-739-907-20	Sequence 20, Appli
38	64.6	2.4	263744	12	US-10-229-834A-6	Sequence 6, Appli
39	64.4	2.4	1156	14	US-10-188-012-34	Sequence 34, Appli
40	64.4	2.4	1156	14	US-10-188-012-36	Sequence 36, Appli
41	64.4	2.4	1282	11	US-09-813-153-68	Sequence 68, Appli
42	64.4	2.4	1380	14	US-10-037-270-59	Sequence 59, Appli
43	63	2.3	382	11	US-09-232-785-348	Sequence 348, App
44	62	2.3	361	11	US-09-232-785-265	Sequence 265, App
45	61.4	2.3	3318	12	US-10-027-632-113926	Sequence 113926,

#### ALIGNMENTS

RESULT 1  
US-10-252-131-8  
Sequence 8, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
Applicant: Levinson, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edwards  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-Seq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Cortuzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864

	TELEX: 66141 PENNIE
:	INFORMATION FOR SEQ ID NO: 8:
:	SEQUENCE CHARACTERISTICS:
:	LENGTH: 2710 base pairs
:	TYPE: nucleic acid
:	STRANDEDNESS: single
:	TOPOLOGY: unknown
:	MOLECULE TYPE: DNA
:	FEATURE:
:	NAME/KEY: CDS
:	LOCATION: 40..885
:	SEQUENCE DESCRIPTION: SEQ ID NO: 8:
:	US-10-252-131-8
Query Match	99.9%; Score 2706.6; DB 12; Length 2710;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2709; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	2 GTGACCCAGCGGTCCGGATTTCCTCCCTCCAGTACTCATATGTTTCAGGCTTACCCTCA 61
Db	
	2 GTGACCCAGCGGTCCGGATTTCCTCCCTCCAGTACTCATATGTTTCAGGCTTACCCTCA 61
QY	62 ACTGTGCTCCTGCTGCTGCATCACTACTTCTTGCAAGCTCATTTGAAGATGTTTAAG 121
Db	62 ACTGTGCTCCTGCTGCTGCATCACTACTTCTTGCAAGCTCATTTGAAGATGTTTAAG 121
QY	122 TTGAGGTTGTAAATAATGCTTATCTGCCCTGCAGATTACA CTCTACCTCATCTTGGACAC 181
Db	122 TTGAGGTTGTAAAATGCTTATCTGCCCTGCAGATTACA CTCTACCTCATCTTGGACAC 181
QY	182 TTGTGCCCTATGTGCTGCGGCGAAGGATTCCTGTCTTGTCACAGTGTACCATGAGTCC 241
Db	182 TTGTGCCCTATGTGCTGCGGCGAAGGATTCCTGTCTTGTCACAGTGTACCATGAGTGC 241
QY	242 TCAGAATCTGATAAAAAATGTGACATATCAGAAAATCCAGAGATCCAGCTTAAAGGGCG 301
Db	242 TCAGAACTGATAAAAAATGTGACATATCAGAAAATCCAGAGATCCAGCTTAAAGGGCG 301
QY	302 ATCTCAACAAAGAGATGTGTCTGTGATATAAAGATGTGA CTCTGATGACCATGGGA 361
Db	302 ATCTCAACAAAGAGATGTGTCTGTGATATAAAGATGTGA CTCTGATGACCATGGGA 361
QY	362 CCTATCTGCTGCAGGATACAGTTCCCTGGTCTTATGAATGATATTAATAATTGAATCTGAAAT 421
Db	362 CCTATCTGCTGCAGGATACAGTTCCCTGGTCTTATGAATGATATTAATAATTGAATCTGAAAT 421
QY	422 TAGACATCAAAGCAGCGCAAGGTCACTCCAGCTCAGACTGCCCATGGGAGCTCTACTACAG 481
Db	422 TAGACATCAAAGCAGCGCAAGGTCACTCCAGCTCAGACTGCCCATGGGAGCTCTACTACAG 481
QY	482 CTYTCTCCAAGAACCTTAAACCAAGGAGAGAAATGTTTCAAGACA CAGACACTGTGTGACCC 541
Db	482 CTYTCTCCAAGAACCTTAAACCAAGGAGAGAAATGTTTCAAGACA CAGACACTGTGTGACCC 541
QY	542 TCCATPATATCAATGGAACAAAAATTTCCACATGGGCTATGAAATTTAAGACTCTGGAG 601
Db	542 TCCATPATATCAATGGAACAAAAATTTCCACATGGGCTATGAAATTTAAGACTCTGGAG 601
QY	602 AAACGATCGAAGCTGTATCCA CATTTGAGTGGGAGTCTCTGGGTGTGAACCTGGGAC 661
Db	602 AAACGATCGAAGCTGTATCCA CATTTGAGTGGGAGTCTCTGGGTGTGAACCTGGGAC 661
QY	662 TTATCATTTGGTGTCTTAAATCCTTAAATGATATTCCTGTAAAGAAAAAGATATGAGATT 721
Db	662 TTATCATTTGGTGTCTTAAATCCTTAAATGATATTCCTGTAAAGAAAAAGATATGAGATT 721
QY	722 TGAGCCTTATTA CACTGGCCAACTTGCTTCAGAGAGGTTGGCAATTCAGAGACAGTCA 781
Db	722 TGAGCCTTATTA CACTGGCCAACTTGCTTCAGAGAGGTTGGCAATTCAGAGACAGTCA 781
QY	782 GGATTTGCTCTGAGGAAAAATATCTTACACATGAGAGAAACGTTATTTGAATGAGAGAAAT 841
Db	782 GGATTTGCTCTGAGGAAAAATATCTTACACATGAGAGAAACGTTATTTGAATGAGAGAAAT 841

QY	842	AAATGAGTCTAATGCTAAGTCAACGACGACGACATCTGACCGCTCTGGACCTGCC	901
Db	842	CAATGAGTACTGCTAGCTGACGACGACGACGACCTGACCGCTCTGGACCTGCC	901
QY	902	ACTTTAAAGGCTGCGCTCAATTTCTGACTTTGGTATTTCCCTTXXTGAAAACATATG	961
Db	902	ACTTTAAAGGCTGCGCTCAATTTCTGACTTTGGTATTTCCCTTXXTGAAAACATATG	961
QY	962	ATATGTCACTTGGGAACCTCATTTGGAGTTCTGACCAAGCAACTGAGAAAAGTTCCA	1021
Db	962	ATATGTCACTTGGGAACCTCATTTGGAGTTCTGACCAAGCAACTGAGAAAAGTTCCA	1021
QY	1022	GTTTTCTGGGATTAATTAACCAAGGGGATTCGACTGTAACTCATGCTACATTTGAAT	1081
Db	1022	GTTTTCTGGGATTAATTAACCAAGGGGATTCGACTGTAACTCATGCTACATTTGAAT	1081
QY	1082	GCTCCATTTTATCCCTGAGTTTCAGGAGTGGATCTCCACTCCAGACCTTCAATCATG	1141
Db	1082	GCTCCATTTTATCCCTGAGTTTCAGGAGTGGATCTCCACTCCAGACCTTCAATCATG	1141
QY	1142	CGTGTGAAGCTCACTGCTGCTTACATACATTTAGAAATGGTTAGTGTGATCTTTGAGA	1201
Db	1142	CGTGTGAAGCTCACTGCTGCTTACATACATTTAGAAATGGTTAGTGTGATCTTTGAGA	1201
QY	1202	CATAGAGTTTGTGATATCCGCAAAAGCTCTTGAAACAGTGGGGGAAATTAAGGGCTAA	1261
Db	1202	CATAGAGTTTGTGATATCCGCAAAAGCTCTTGAAACAGTGGGGGAAATTAAGGGCTAA	1261
QY	1262	GATAGGAAGTGGCGCTCTTTGTTGATTTGAGAAAATCTTAAAGATTTGATGCTTTTCT	1321
Db	1262	GATAGGAAGTGGCGCTCTTTGTTGATTTGAGAAAATCTTAAAGATTTGATGCTTTTCT	1321
QY	1322	AGAAATTTCTGACCTTGAAAGATTAAGAAAAGCCAGGTGGCATATCTTAAACGATAT	1381
Db	1322	AGAAATTTCTGACCTTGAAAGATTAAGAAAAGCCAGGTGGCATATCTTAAACGATAT	1381
QY	1382	AACTTGGGAACCTTAGCGAGGAGGTATTAATCAAGGTCAAGCAGGGCTATCTGGTAA	1441
Db	1382	AACTTGGGAACCTTAGCGAGGAGGTATTAATCAAGGTCAAGCAGGGCTATCTGGTAA	1441
QY	1442	AGACTGTCTCMCACTCCAAAGACGAAATTAACATAGACAGCAGGAGGCTGGAGATGA	1501
Db	1442	AGACTGTCTCMCACTCCAAAGACGAAATTAACATAGACAGCAGGAGGCTGGAGATGA	1501
QY	1502	GGCTTCGACAGTGAAGGTGATTTGTATCAAGCAAGAGAAATCTATATTTGATCTGAC	1561
Db	1502	GGCTTCGACAGTGAAGGTGATTTGTATCAAGCAAGAGAAATCTATATTTGATCTGAC	1561
QY	1562	CCAGATGAAAAAGCTAGGCGGTGTAAGACATGCTTTGATGACTCAAGAGATGGAGAGTAA	1621
Db	1562	CCAGATGAAAAAGCTAGGCGGTGTAAGACATGCTTTGATGACTCAAGAGATGGAGAGTAA	1621
QY	1622	AGGCAACAACAGATCCCGGGGCTTGCCTGACAGTCAAGTCAAGTGTGCTGAGTTCCAA	1681
Db	1622	AGGCAACAACAGATCCCGGGGCTTGCCTGACAGTCAAGTCAAGTGTGCTGAGTTCCAA	1681
QY	1682	GTCCAACAAGATCCCTGTCTCAAGATTAAGATGACTGATATCTGGGCGCATATCTGAGG	1741
Db	1682	GTCCAACAAGATCCCTGTCTCAAGATTAAGATGACTGATATCTGGGCGCATATCTGAGG	1741
QY	1742	GGTGTGCTCTCGCTCAGAGAGACATGCAATGACCTTGCACACACACACACACAC	1801
Db	1742	GGTGTGCTCTCGCTCAGAGAGACATGCAATGACCTTGCACACACACACACACAC	1801
QY	1802	ACACACACACACACACACACACATGAAATGAAGTTCTCTCTGTGCTGCTAAC	1861
Db	1802	ACACACACACACACACACACACATGAAATGAAGTTCTCTCTGTGCTGCTAAC	1861
QY	1862	TCTCTAATAACATGATCTCTACAGACTCTCTGCTCTGTTAAGACATGAATGGAG	1921
Db	1862	TCTCTAATAACATGATCTCTACAGACTCTCTGCTCTGTTAAGACATGAATGGAG	1921

QY 1922 CATGGCAGAGCAGTCCAGTAATTTATTCAGACCTCAGAGGCTGGAGCAGAGCTGGA 1981  
 DB 1922 CATGGCAGAGCAGTCCAGTAATTTATTCAGACCTCAGAGGCTGGAGCAGAGCTGGA 1981  
 QY 1982 GAGTTCAGAGCAGTGTGCTCCCACTGCGCAGACTCTTTCTTACACAGAAAAAGTTACC 2041  
 DB 1982 GAGTTCAGAGCAGTGTGCTCCCACTGCGCAGACTCTTTCTTACACAGAAAAAGTTACC 2041  
 QY 2042 CGCAGACAGCTGTGCTGTGTAAGGAAACCTGCGAAAGCAACCTTGACTGTGTC 2101  
 DB 2042 CGCAGACAGCTGTGCTGTGTAAGGAAACCTGCGAAAGCAACCTTGACTGTGTC 2101  
 QY 2102 TGCTCAAGGGGAACTGCACTCAGACCACTTCTCCATTCCTGAGGAAACTGAGCTGTTC 2161  
 DB 2102 TGCTCAAGGGGAACTGCACTCAGACCACTTCTCCATTCCTGAGGAAACTGAGCTGTTC 2161  
 QY 2162 TGACGAGAGAACACCGGTGACTGGGACATAGAGGCGAGAGCTCTTGCAGCAATCTATA 2221  
 DB 2162 TGACGAGAGAACACCGGTGACTGGGACATAGAGGCGAGAGCTCTTGCAGCAATCTATA 2221  
 QY 2222 TAGTCAGCAAAATATCTTTGGAGGACAGTGTGTCACCAATTTGATTCAGACCGGTG 2281  
 DB 2222 TAGTCAGCAAAATATCTTTGGAGGACAGTGTGTCACCAATTTGATTCAGACCGGTG 2281  
 QY 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGCCTTGAATCTGTGCTGCTGCC 2341  
 DB 2282 ACCTCAGTTTCACTGCTTACAGCTGCTGCGCAGTGCCTTGAATCTGTGCTGCTGCC 2341  
 QY 2342 ATCTATACAGAAATCAATTAATTAAGACCCCGAGTGAATAATTAAGTGAAGAAAGGT 2401  
 DB 2342 ATCTATACAGAAATCAATTAATTAAGACCCCGAGTGAATAATTAAGTGAAGAAAGGT 2401  
 QY 2402 AGCTTGTTCAGAAATTTTTTGGAGGACAGTGTGTCACCAATTTGATTCAGACCGGTG 2461  
 DB 2402 AGCTTGTTCAGAAATTTTTTGGAGGACAGTGTGTCACCAATTTGATTCAGACCGGTG 2461  
 QY 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTCATGATGAAATTTGTTGTTAGGT 2521  
 DB 2462 TAGTGAAGACACCAAAACCTGTGTACCGTTTTCATGATGAAATTTGTTGTTAGGT 2521  
 QY 2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTGAAGTGTGATGAGAGGAGCAATC 2581  
 DB 2522 TGCTTCTAGCTAGCTGTGAGGTCTGTGCTTTCTTGAAGTGTGATGAGAGGAGCAATC 2581  
 QY 2582 TTAACAAATTCATTAAGATTAACAGCTTCATGACAGAGGAAACCTAATCTCAATGTT 2641  
 DB 2582 TTAACAAATTCATTAAGATTAACAGCTTCATGACAGAGGAAACCTAATCTCAATGTT 2641  
 QY 2642 TTAAGTATATAAACTGTACTGCAAAAGTCTTGTGATGATGAAATTTGTTGTTAGGT 2701  
 DB 2642 TTAAGTATATAAACTGTACTGCAAAAGTCTTGTGATGATGAAATTTGTTGTTAGGT 2701  
 QY 2702 GGGGGGGC 2710  
 DB 2702 GGGGGGGC 2710

RESULT 2  
 US-10-004-633-8  
 ; Sequence 8, Application US/10004633  
 ; Publication No. US20030069196A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; APPLICANT: Lloyd, Clare M.  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
 ; FILE REFERENCE: 7853-125  
 ; CURRENT APPLICATION NUMBER: US/10/004,633  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

; PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patencin Ver. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 2710  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: modified base  
 ; LOCATION: all "n" positions  
 ; OTHER INFORMATION: n=a, c, g, or t  
 US-10-004-633-8

Query Match 99.9%; Score 2706.6; DB 14; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 GTGACCCACGCGTCCGATTTCCCTCCCAAGTACTAGTTTCAAGCTTTACCTCA 61  
 DB 2 GTGACCCACGCGTCCGATTTCCCTCCCAAGTACTAGTTTCAAGCTTTACCTCA 61  
 QY 62 ACTGTCCTGCTGCTGCTGCAACTACTTGCAGAGGCAATTTGAAGTGTATAAG 121  
 DB 62 ACTGTCCTGCTGCTGCTGCAACTACTTGCAGAGGCAATTTGAAGTGTATAAG 121  
 QY 122 TTGAGTGTGTAAATGCTTATCTGCTGAGTTACACTTACTACTCTGAGAC 181  
 DB 122 TTGAGTGTGTAAATGCTTATCTGCTGAGTTACACTTACTACTCTGAGAC 181  
 QY 182 TTGTCCTATGCTGTGCGGCAAGGATTTCTGCTTGTGACAGTGTACCAATGAGTTC 241  
 DB 182 TTGTCCTATGCTGTGCGGCAAGGATTTCTGCTTGTGACAGTGTACCAATGAGTTC 241  
 QY 242 TCAGAACTGATGAAGAAATGTGATATATCAAGATTCAGAGATTCAGCTTAAAGGGG 301  
 DB 242 TCAGAACTGATGAAGAAATGTGATATATCAAGATTCAGAGATTCAGCTTAAAGGGG 301  
 QY 302 ATCTCAACAAAGAGATGTCTGTATCATTAAGATGTGACTGTGATGACATGGA 361  
 DB 302 ATCTCAACAAAGAGATGTCTGTATCATTAAGATGTGACTGTGATGACATGGA 361  
 QY 362 CCTACTGCTGAGATACAGTTCCCTGTCTTATGATGATTAATAATTAAGTGAAT 421  
 DB 362 CCTACTGCTGAGATACAGTTCCCTGTCTTATGATGATTAATAATTAAGTGAAT 421  
 QY 422 TAGACATCAAGACAGGCAAGGCTCACTCCAGCTCAAGCTGCGCATGGAGCTTACTACG 481  
 DB 422 TAGACATCAAGACAGGCAAGGCTCACTCCAGCTCAAGCTGCGCATGGAGCTTACTACG 481  
 QY 482 CTTCTCAAGAACTTAACACAGAGAAATGTGTTACAGACAGACACTGTGACCC 541  
 DB 482 CTTCTCAAGAACTTAACACAGAGAAATGTGTTACAGACAGACACTGTGACCC 541  
 QY 542 TCCATATTAACATGAAACAAATTTTCCATGAGGCTGATGAAATTAAGACTGAG 601  
 DB 542 TCCATATTAACATGAAACAAATTTTCCATGAGGCTGATGAAATTAAGACTGAG 601  
 QY 602 AAACGATCAAGACGCTATCCACATTTGAGAGGAGTCTGCTGAGGCTGAGCCCTGAC 661  
 DB 602 AAACGATCAAGACGCTATCCACATTTGAGAGGAGTCTGCTGAGGCTGAGCCCTGAC 661  
 QY 662 TTATCATGTTGTCTTAATCTTAATGATGATTCCTGTAAAGAAAGAGTATGAGTT 721  
 DB 662 TTATCATGTTGTCTTAATCTTAATGATGATTCCTGTAAAGAAAGAGTATGAGTT 721  
 QY 722 TGAGCTTATTAACATGAGCACTTGTCTTCAAGAGGCTTGAATTCAGAGCAGTCA 781  
 DB 722 TGAGCTTATTAACATGAGCACTTGTCTTCAAGAGGCTTGAATTCAGAGCAGTCA 781

Oy	782	GGATTGCGCTGAGGAAAAATATCTACACATGAGAGAGAAAGTAATAGAGGGAATT	841
Db	782	GGATTGCGCTCTGAGGAAAAATATCTACACATGAGAGAGAAAGTAATAGAGGGAATT	841
Oy	842	CAAAATGAGTACTACTGCTACGTCAAAGCCAGACAGCACTCTGACCGGCTCTGACTGCC	901
Db	842	CAAAATGAGTACTACTGCTACGTCAAAGCCAGACAGCACTCTGACCGGCTCTGACTGCC	901
Oy	902	ACTTTTAAAGGCTCGGCTTCAATTTCTGACTTTGGTAATTTCCCTTTKTGAAAACTATGTG	961
Db	902	ACTTTTAAAGGCTCGGCTTCAATTTCTGACTTTGGTAATTTCCCTTTKTGAAAACTATGTG	961
Oy	962	AATATGCACTTGGCAACCTCATTGGAGGTTCTGACACAGCCACTGAGAAAAAGGTTCCA	1021
Db	962	AATATGCACTTGGCAACCTCATTGGAGGTTCTGACACAGCCACTGAGAAAAAGGTTCCA	1021
Oy	1022	GTTTTCTGGGGAATTAATTACTCACAAGGGGATTCGACTGTAACTATGCTACATTGAAT	1081
Db	1022	GTTTTCTGGGGAATTAATTACTCACAAGGGGATTCGACTGTAACTATGCTACATTGAAT	1081
Oy	1082	GCTCCATTTTATCCTGAGTTTCAAGGATCGATCTCCACTCCAGAGCTTCAATCATG	1141
Db	1082	GCTCCATTTTATCCTGAGTTTCAAGGATCGATCTCCACTCCAGAGCTTCAATCATG	1141
Oy	1142	CGTTTGAAGGTCACGTCGTCCTTCAATCATTAGAAAGGTATGTAATGCTTTGAGA	1201
Db	1142	CGTTTGAAGGTCACGTCGTCCTTCAATCATTAGAAAGGTATGTAATGCTTTGAGA	1201
Oy	1202	CATAGAGGTTTGTGATATCCGCAAGGCTCTGAAAGGTATGAGGGAATTAAGGGCTTA	1261
Db	1202	CATAGAGGTTTGTGATATCCGCAAGGCTCTGAAAGGTATGAGGGAATTAAGGGCTTA	1261
Oy	1262	GATAGAGAGTGCGGCTTTTGTGATGTTGAAAACTTTAAAGAGTTGGTAGCTTTCT	1321
Db	1262	GATAGAGAGTGCGGCTTTTGTGATGTTGAAAACTTTAAAGAGTTGGTAGCTTTCT	1321
Oy	1322	AAGAGATTTCTGACCTTGAAGAATTAAGAAAAAGCCAGGTGCGATATGCTTAAACAGATAT	1381
Db	1322	AAGAGATTTCTGACCTTGAAGAATTAAGAAAAAGCCAGGTGCGATATGCTTAAACAGATAT	1381
Oy	1382	AACCTTGGGAACCTTAGAGGCGAGAGGTGATTAAGGTCAGCCAGGGCTATGCTGCTA	1441
Db	1382	AACCTTGGGAACCTTAGAGGCGAGAGGTGATTAAGGTCAGGTCAGCGGGCTATGCTGCTA	1441
Oy	1442	AGACTGTCTCAGCATCCAAAGACGAAATTAACATAGAGACAGCAGAGGCTGAGATGA	1501
Db	1442	AGACTGTCTCAGCATCCAAAGACGAAATTAACATAGAGACAGCAGAGGCTGAGATGA	1501
Oy	1502	GGCTCGGACAGTGAAGTGATTTGTGTACAGACAGAGGAATTAATTTTGAATCGTAGACC	1561
Db	1502	GGCTCGGACAGTGAAGTGATTTGTGTGTACAGAGCAGAGGAATTAATTTTGAATCGTAGACC	1561
Oy	1562	CCACATGAAGAAAGCTTAGGCTGTGTAGAGATCTTTGTAGACTTCAAGATGAGAGGTTAA	1621
Db	1562	CCACATGAAGAAAGCTTAGGCTGTGTGTAGAGATCTTTGTAGACTTCAAGATGAGAGGTTAA	1621
Oy	1622	AGGCAACAACAGATCCCGGGGCTTGTGTCAGCTAGCTTAGCTTAGCTGAGTTCCAA	1681
Db	1622	AGGCAACAACAGATCCCGGGGCTTGTGTCAGCTAGCTTAGCTTAGCTGAGTTCCAA	1681
Oy	1682	GTTCCACAGAGATGTCCTGTCTCAGAGTAAGATGAGTATCTGGCCATGTCATGAGG	1741
Db	1682	GTTCCACAGAGATGTCCTGTCTCAGAGTAAGATGAGTATCTGGCCATGTCATGAGG	1741
Oy	1742	GGTTGTCCTCTCTCTCAGAGAGATGATCAATGCCCTGACACACACACACACAC	1801
Db	1742	GGTTGTCCTCTCTCTCAGAGAGATGATCAATGCCCTGACACACACACACACACAC	1801
Oy	1802	ACACACACACACACACACACACACACACATGATGAAGGTTCTCTCTGTGCTGTACC	1861
Db	1802	ACACACACACACACACACACACACACACATGATGAAGGTTCTCTCTGTGCTGTACC	1861
Oy	1862	TCTCTATTAACATGATATCTCTACAGGACTCTCCTCTGTGTAAACATGATGGAG	1921

Db	1662	TCTCTATTAACAATGATATCTTAACAGACCTCTCTCTGCTCTGTTAAACATGATGGGAG	1921
QY	1922	CATGGCAGAGCAGTCCAGTAATTTATTTCCAGCACTCAAGAGGCTGGAGCAGACGTTGA	1981
Db	1922	CATGGCAGAGCAGTCCAGTAATTTATTTCCAGCACTCAAGAGGCTGGAGCAGACGTTGA	1981
QY	1982	GAGTTTCAGSACATGTGGCCCAACACATGCGCAGACTCTTTCTTAACAAGAAAAGGTTACC	2041
Db	1982	GAGTTTCAGSACATGTGGCCCAACACATGCGCAGACTCTTTCTTAACAAGAAAAGGTTACC	2041
QY	2042	CGCAAGCAGCCTGCTGCTGTAAAAAGGAAAACCCGCGAAAAGCCAACTTTGACTGTTGTG	2101
Db	2042	CGCAAGCAGCCTGCTGCTGTAAAAAGGAAAACCCGCGAAAAGCCAACTTTGACTGTTGTG	2101
QY	2102	TGCTCAAGGGGAACTGACTCAGACCAACTTCTCCATTTCTTGGAGGAAACTGGAGCTGTTTC	2161
Db	2102	TGCTCAAGGGGAACTGACTCAGACCAACTTCTCCATTTCTTGGAGGAAACTGGAGCTGTTTC	2161
QY	2162	TGACAGAAAGAACCAACCGGTGACTGGGACATACAGAGGCAAGCTCTTGACCAATCTATA	2221
Db	2162	TGACAGAAAGAACCAACCGGTGACTGGGACATACAGAGGCAAGCTCTTGACCAATCTATA	2221
QY	2222	TAGTCAGCAAAATATTCTTTGGGAGGAGCAGTCGACCAAAATTGATTTCCAAAGCGGTGG	2281
Db	2222	TAGTCAGCAAAATATTCTTTGGGAGGAGCAGTCGACCAAAATTGATTTCCAAAGCGGTGG	2281
QY	2282	ACCTCAGTTTCATCTGGCTTACAGCTGCTGTCGCCAGTGCCTTGATCTGTGCTGGCTCC	2341
Db	2282	ACCTCAGTTTCATCTGGCTTACAGCTGCTGTCGCCAGTGCCTTGATCTGTGCTGGCTCC	2341
QY	2342	ATCTATAACAGAAATCAAAATTAATATGACCCGAGATGAAAAATTTAAGTAGACAAAAGT	2401
Db	2342	ATCTATAACAGAAATCAAAATTAATATGACCCGAGATGAAAAATTTAAGTAGACAAAAGT	2401
QY	2402	AGCTTTGTTCAAAATTTTTTTTGCATTTGGGAGGCAACGTGTATCATCAGAGGACATCTGT	2461
Db	2402	AGCTTTGTTCAAAATTTTTTTTGCATTTGGGAGGCAACGTGTATCATCAGAGGACATCTGT	2461
QY	2462	TAGTAGAGACACAAAAACCTGTGTACCGTTTTTTCATGATGAATTTGTTGTTTAAAGT	2521
Db	2462	TAGTAGAGACACAAAAACCTGTGTACCGTTTTTTCATGATGAATTTGTTGTTTAAAGT	2521
QY	2522	TGCTTTCTAGCTAGCTGTGGAGAGTCTGTGCTTTCTTTAGGTGGGTATGGAAGGAAACATC	2581
Db	2522	TGCTTTCTAGCTAGCTGTGGAGAGTCTGTGCTTTCTTTAGGTGGGTATGGAAGGAAACATC	2581
QY	2582	TAACAAAATCCATTAGAGATATACAGCTCTCATGAGAAAGGAAAACTAATCTCAAAATGTT	2641
Db	2582	TAACAAAATCCATTAGAGATATACAGCTCTCATGAGAAAGGAAAACTAATCTCAAAATGTT	2641
QY	2642	TTAAAGTATATAAATCTGACTGCGAAAGTACTTTGACATTAATAAAAAAAAAAAAAAAG	2701
Db	2642	TTAAAGTATATAAATCTGACTGCGAAAGTACTTTGACATTAATAAAAAAAAAAAAAAAG	2701
QY	2702	GCGCGCGCGC 2710	
Db	2702	GCGCGCGCGC 2710	
RESULT 3			
US-10-188-012-10			
; Sequence 10, Application US/10188012			
; Publication No. US20030124114A1			
GENERAL INFORMATION:			
APPLICANT: McInerite, Jennifer Jones			
APPLICANT: Umetani, Dale T.			
APPLICANT: Dekruyf, Rosemarie			
APPLICANT: Kuchroo, Vijay			
APPLICANT: Freeman, Gordon J.			
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of			
FILE REFERENCE: STDAN-235			



; CURRENT APPLICATION NUMBER: US/10/188,012  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 2725  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-188-012-10

Query Match 96.8%; Score 2624.6; DB 14; Length 2725;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2655; Conservative 6; Mismatches 20; Indels 2; Gaps 2;

QY 18 GGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTTACCTCAACTGTGCTGCTGCT 77  
DB 45 GGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTTACCTCAACTGTGCTGCTGCT 104  
QY 78 GCTGCACTACTACTGCAAGGTCATTTGGAAGAAGGTTATTAAGGTTGAGTTGTAATAA 137  
DB 105 GCTGCACTACTACTGCAAGGTCATTTGGAAGAAGGTTATTAAGGTTGAGTTGTAATAA 164  
QY 138 TGCCATATGCTGCTGCAAGTACTCATGTTTTCAGGCTTTACCTCAACTGTGCTGCTG 197  
DB 165 TGCCATATGCTGCTGCAAGTACTCATGTTTTCAGGCTTTACCTCAACTGTGCTGCTG 224  
QY 198 GGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257  
DB 225 GGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 284  
QY 258 AAATGTGACATATGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCA 317  
DB 285 AAATGTGACATATGCAAGGATTTGCAAGGATTTGCAAGGATTTGCAAGGATTTGCA 344  
QY 318 TGTGCTCTGATCATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 377  
DB 345 TGTGCTCTGATCATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 404  
QY 378 ACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437  
DB 405 ACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 464  
QY 438 CAAGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497  
DB 465 CAAGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 524  
QY 498 AACCAAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557  
DB 525 AACCAAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584  
QY 558 AACCAAAATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617  
DB 585 AACCAAAATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644  
QY 618 TATCAACATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677  
DB 645 TATCAACATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 704  
QY 678 AATCTTAATGATGATTTCTGTAAGAAAAGAAAGTATATGAGTTGAGCTTATTAAC 737  
DB 705 AATCTTAATGATGATTTCTGTAAGAAAAGAAAGTATATGAGTTGAGCTTATTAAC 764  
QY 738 GGGCAACTTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797  
DB 765 GGGCAACTTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824  
QY 798 AAATATCTAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857  
DB 825 AAATATCTAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 884  
QY 858 CTACGTCAACAGCCAGAGCCATCTGACCGGCTCTGACCTGCACTTTTAAGGCTGCG 917

DB 885 CTACGTCAACAGCCAGAGCCATCTGACCGGCTCTGACCTTGAAGCTGCG 944  
QY 918 CTTCATTTCTGACTTTGGTATTTTCCCTTTTGTGGAAGATATGATGATGATGATG 977  
DB 945 CTTCATTTCTGACTTTGGTATTTTCCCTTTTGTGGAAGATATGATGATGATGATG 1003  
QY 978 CCTCATTTGAGGATTTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037  
DB 1004 CCTCATTTGAGGATTTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1063  
QY 1038 TAACTCAAGGAGGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097  
DB 1064 TAACTCAAGGAGGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1123  
QY 1098 GAGTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1157  
DB 1124 GAGTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1188  
QY 1158 GGTGCTTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1217  
DB 1184 GGTGCTTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1243  
QY 1218 ATATCCGCAAGGCTCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1277  
DB 1244 ATATCCGCAAGGCTCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1303  
QY 1278 CTGTTGATGATTTGGAAGATTTTAAAGATTTGAGGATTTTAAAGATTTGAG 1337  
DB 1304 CTGTTGATGATTTGGAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1362  
QY 1338 GAAAGATTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1397  
DB 1363 GAAAGATTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1422  
QY 1398 GCAGGAGGATTAAGTTCAAGGATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1457  
DB 1423 GCAGGAGGATTAAGTTCAAGGATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1482  
QY 1458 CAAGACAAATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1517  
DB 1483 CAAGACAAATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1542  
QY 1518 TGCAATTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1577  
DB 1543 TGCAATTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 160  
QY 1578 GGCCTGTAGAGCATGCTGTAAGATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1637  
DB 1603 GGCCTGTAGAGCATGCTGTAAGATTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1662  
QY 1638 CGGGCTTGGCTGCAAGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCT 1697  
DB 1663 CGGGCTTGGCTGCAAGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCT 1722  
QY 1698 GTCTCAAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1757  
DB 1723 GTCTCAAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1782  
QY 1758 CAGAAGAGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 1817  
DB 1783 CAGAAGAGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 1842  
QY 1818 ACACACACACATGAAATGAAAGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1877  
DB 1843 ACACACACACATGAAATGAAAGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1902  
QY 1878 CTCTACAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1937  
DB 1903 CTCTACAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1962  
QY 1938 AGTAATTTATTCAGACCTCAGAAAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 1997

Db 1963 AGTAATTATTCAGACACTCAGAAAGCGCTGAGACAAAGCTGAGAGACTTCAGAGACACTG 2022  
QY 1998 TGGCCCAACACTGCGACACTCTTCTTAACAAGAAAAAGTTACCCGCAAGACACTGCTG 2057  
Db 2023 TGGCCCAACACTGCGACACTCTTCTTAACAAGAAAAAGTTACCCGCAAGACACTGCTG 2082  
QY 2058 TCTGTAAAGGAAACCTGCGAAGGCAAACTTGTGCTGTGTGCTGCAAGGGAACCTG 2117  
Db 2083 TCTGTAAAGGAAACCTGCGAAGGCAAACTTGTGCTGTGTGCTGCAAGGGAACCTG 2142  
QY 2118 ACTCAGACAACTTCTCACTTCTGAGAGAACTGAGCTGTTCTGACAGAAAGAACACC 2177  
Db 2143 ACTCAGACAACTTCTCACTTCTGAGAGAACTGAGCTGTTCTGACAGAAAGAACACC 2202  
QY 2178 GGTGACTGGGACATACGAAAGGACAGCTCTTGACAGCAATCTATATAGTCAGCAAAATATT 2237  
Db 2203 GGTGACTGGGACATACGAAAGGACAGCTCTTGACAGCAATCTATATAGTCAGCAAAATATT 2262  
QY 2238 CTTTGGGAGGACAGTGTGTCACCAATTTGATTTCCAAAGCGGTGAGCTGCTGCTCAATCTG 2297  
Db 2263 CTTTGGGAGGACAGTGTGTCACCAATTTGATTTCCAAAGCGGTGAGCTGCTGCTCAATCTG 2322  
QY 2298 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2357  
Db 2323 GCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2382  
QY 2358 AATTAAATAGACCCCGAGTGAATAATTTAAAGTACGAAAGGAGTACCTTGTTCGAAGAT 2417  
Db 2383 AATTAAATAGACCCCGAGTGAATAATTTAAAGTACGAAAGGAGTACCTTGTTCGAAGAT 2442  
QY 2418 TTTTGTGCAATGGGAGGACACTGTATCATCAGAGGACATCTGTTAGTGGAGACACCAAA 2477  
Db 2443 TTTTGTGCAATGGGAGGACACTGTATCATCAGAGGACATCTGTTAGTGGAGACACCAAA 2502  
QY 2478 ACCGTGTGACCGTTTTTTCATGATGATGATTTTGTGTTAGTGGTCTTCTAGCTAGCTG 2537  
Db 2503 ACCGTGTGACCGTTTTTTCATGATGATGATTTTGTGTTAGTGGTCTTCTAGCTAGCTG 2562  
QY 2538 TGGAGTCTGCTGCTTCTTAAAGTGGTATGAAAGGAGACCAATCTTAAAGAAATCCATTAG 2597  
Db 2563 TGGAGTCTGCTGCTTCTTAAAGTGGTATGAAAGGAGACCAATCTTAAAGAAATCCATTAG 2622  
QY 2598 AGATAACAGCTCTCATGACAGAGGAGAACTAATCTCAATGTTTTTAAATTAATAACT 2657  
Db 2623 AGATAACAGCTCTCATGACAGAGGAGAACTAATCTCAATGTTTTTAAATTAATAACT 2682  
QY 2658 GTAATGCAAAAGTACTTGAAGCAATAAAAAA 2700  
Db 2683 GTAATGCAAAAGTACTTGAAGCAATAAAAAA 2725

RESULT 4  
US-10-188-012-12  
; Sequence 12, Application US/10188012  
; Publication No. US20030124114A1  
; GENERAL INFORMATION:  
; APPLICANT: McInchre, Jennifer Jones  
; APPLICANT: Umetsu, Dale T.  
; APPLICANT: Dekruyf, Rosemarie  
; APPLICANT: Kuchroo, Vijay  
; APPLICANT: Freeman, Gordon J.  
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
; FILE REFERENCE: STAN-235  
; CURRENT APPLICATION NUMBER: US/10/188,012  
; PRIORITY FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: 60/302,344  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 862  
; TYPE: DNA

; ORGANISM: Mus musculus  
US-10-188-012-12  
Query Match 31.2%; Score 844.4; DB 14; length 862;  
Best Local Similarity 98.7%; Pred. No. 2,1e-245;  
Matches 851; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 24 CCCCTCCCAAGTACATCTTTTCAAGTCTTACCTCTCAACTGTGTCTGTCTGTGCA 83  
Db 1 CCCCTCCCAAGTACATCTTTTCAAGTCTTACCTCTCAACTGTGTCTGTCTGTGCA 60  
QY 84 ACTACTACTTGGCAAGGTCATGGAAGATGTTATTAAGTTGAGTTGTTAAATATCCCTA 143  
Db -61 ACTACTACTTGGCAAGGTCATGGAAGATGTTATTAAGTTGAGTTGTTAAATATCCCTA 120  
QY 144 TCTGCTGCTGCACTTACCTTACCTACATCTGGAACACTTGTGCTGCTGCTGCTGCA 203  
Db 121 TCTGCTGCTGCACTTACCTTACCTACATCTGGAACACTTGTGCTGCTGCTGCTGCA 180  
QY 204 GGGATTTCTGCTTGTGTCACTGTACCAATGATGCTGCAAGATGATGATGATGATGAT 263  
Db 181 GGGATTTCTGCTTGTGTCACTGTACCAATGATGCTGCAAGATGATGATGATGATGAT 240  
QY 264 GACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCTCAACAAAGAGATGTGTC 323  
Db 241 GACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCTCAACAAAGAGAGATGTGTC 300  
QY 324 TCTGATCATTAAGATGATGATCTGATGATGATGATGATGATGATGATGATGATGAT 383  
Db 301 TCTGATCATTAAGATGATGATCTGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 384 CCTGTGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443  
Db 361 CCTGTGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 444 CACTCAGCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503  
Db 421 CACTCAGCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 504 GAGAGAAATGCTTACAGACACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 563  
Db 481 GAGAGAAATGCTTACAGACACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 564 AATTTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623  
Db 541 AATTTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 624 CATTGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683  
Db 601 CATTGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 684 TAAATGTAATTCCTGTGAAGAAAGATTAATGAGTTGAGCTTATTAACCTGCGCA 743  
Db 661 TAAATGTAATTCCTGTGAAGAAAGATTAATGAGTTGAGCTTATTAACCTGCGCA 720  
QY 744 CTTGCTTCCAGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803  
Db 721 CTTGCTTCCAGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 804 CTACACCATGAGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 863  
Db 781 CTACACCATGAGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 864 CAACAGCCAGACGCACTCTGA 885  
Db 841 CAACAGCCAGACGCACTCTGA 862

RESULT 5  
US-10-252-131-22  
; Sequence 22, Application US/10252131  
; Publication No. US20030158399A1  
; GENERAL INFORMATION:

```

1  APPLICANT: Levinson, Douglas A.
2  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
3  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
4  NUMBER OF SEQUENCES: 37
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Pennie & Edmonds
7  STREET: 1155 Avenue of the Americas
8  CITY: New York
9  STATE: New York
10 COUNTRY: USA
11 ZIP: 10036/2711
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: PasteSeq Version 2.0
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/10/252,131
21 FILING DATE: 20-Sep-2002
22 CLASSIFICATION: <Unknown>
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/06/609,563
26 FILING DATE: 01-MAR-1996
27 APPLICATION NUMBER: US 08/487,748
28 FILING DATE: 07-JUN-1995
29 APPLICATION NUMBER: US 08/398,633
30 FILING DATE: 03-MAR-1995
31
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Coruzzi, Laura A.
34 REGISTRATION NUMBER: 30,742
35 REFERENCE/DOCKET NUMBER: 7853-048
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 212-790-9090
38 TELEFAX: 212-869-8864
39
40 TELEX: 66141 PENNIE
41
42 INFORMATION FOR SEQ ID NO: 22:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 843 base pairs
45 TYPE: nucleic acid
46 STRANDEDNESS: single
47 TOPOLOGY: linear
48
49 MOLECULE TYPE: DNA
50
51 SEQUENCE DESCRIPTION: SEQ ID NO: 22:
52
53 US-10-252-131-22
54
55 Query Match 31.1%; Score 843; DB 12; Length 843;
56 Best Local Similarity 100.0%; Pred.No. 5.4e-245; Indels 0; Gaps 0
57 Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921

```

OY	400	ATATAAAAATTAGAACTGAAATTTGACATCAAGCAGCAAGGTCACTCACTCAAGCT	459
Db	361	GATAAAAAATTAGAACTGAAATTTGACATCAAGCAGCAAGGTCACTCACTCAAGCT	420
OY	460	GCCCATGGGGACTCTACTACAGCTTCTCCAGAACCCCTTAACACGAGAGAAATGGTTCA	519
Db	421	GCCCATGGGGACTCTACTACAGCTTCTCCAGAACCCCTTAACACGAGAGAAATGGTTCA	480
OY	520	GAGACACAGACACTGGTGACCCCTCCATATAACATGGAACAAAAATTTCCATATGGGCT	579
Db	481	GAGACACAGACACTGGTGACCCCTCCATATAACATGGAACAAAAATTTCCATATGGGCT	540
OY	580	GATGAAATTTAGAGACTCTGAGAAACGATCGAAATCTGTATCACATTTGGAGTGGAGTTC	639
Db	541	GATGAAATTTAGAGACTCTGAGAAACGATCGAAATCTGTATCACATTTGGAGTGGAGTTC	600
OY	640	TTCTGTGGGTTGACCCCTGGCACTTATCATATGGTGTCTTAATTCCTTAATAGTAATTCCTG	699
Db	601	TTCTGTGGGTTGACCCCTGGCACTTATCATATGGTGTCTTAATTCCTTAATAGTAATTCCTG	660
OY	700	AAGAAAAAGAACTTATCGAGTTTGGACCTTATTAACATGCGCAACTTGGCTCCAGAGGG	759
Db	661	AAGAAAAAGAACTTATCGAGTTTGGACCTTATTAACATGCGCAACTTGGCTCCAGAGGG	720
OY	760	TTGGCAATGCGAGAGCAGTCAGATTGCTCTGAGAAAAATATATCAACAATCGAGAG	819
Db	721	TTGGCAATGCGAGAGCAGTCAGATTGCTCTGAGAAAAATATATCAACAATCGAGAG	780
OY	820	AACGTATATTGAAGTGGAGAAATTCAAATAGATCTACTGCTACGTCAACAGCCAGCAGCCA	879
Db	781	AACGTATATTGAAGTGGAGAAATTCAAATAGATCTACTGCTACGTCAACAGCCAGCAGCCA	840
OY	880	TCC 882	
Db	841	TCC 843	

RESULT 6  
US-10-004-633-22  
Sequence 22, Application US/10004633  
Publication No. US20030069196A1  
GENERAL INFORMATION:  
APPLICANT: Levinson, Douglas A.  
APPLICANT: Lloyd, Clare M.  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF IMMUNE DISORDERS  
FILE REFERENCE: 7853-115  
CURRENT APPLICATION NUMBER: US/10/004,633  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-03-01  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/487,748  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/398,633  
PRIOR FILING DATE: EARLIER FILING DATE: 1995-03-05  
NUMBER OF SEQ. ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 843  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
US-10-004-633-22  
OTHER INFORMATION: Description of Artificial Sequence: primer

Query Match 31.1%; Score 843; DB 14; Length 843;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-245;  
 Matches 843; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0

40 ATGTTTCAGGCTTACCCCTCACTGTGTCTCGCTGCTGCAACTACTACTTGGCAAG 99

Db 1 AAGTTTTCAGGCTTACCTGCACTGTGTCTGCTGCTGCACTACTTACTTGCAGG 60  
QY 100 TCATTGGAAGATGTTAATAAGTTGAGTTGTAATAATGCTATCTGCCCTGCACTTAC 159  
Db 61 TCATTGGAAGATGTTAATAAGTTGAGTTGTAATAATGCTATCTGCCCTGCACTTAC 120  
QY 160 ACTCTACCTACATCTGGAACACTTGTGCTTATGTGCTGAGGCAAGGATTTCTCTTGG 219  
Db 121 ACTCTACCTACATCTGGAACACTTGTGCTTATGTGCTGAGGCAAGGATTTCTCTTGG 180  
QY 220 TCACAGTGTACCAATGATGTGCTCAGAACTGATGAAAGAAATGACATATCAGAAATCC 279  
Db 181 TCACAGTGTACCAATGATGTGCTCAGAACTGATGAAAGAAATGACATATCAGAAATCC 240  
QY 280 AGCAGATACCAAGCTAAAGGGCGATCTCAACAAAGAGATGTCTCTGATCATMAAGAT 339  
Db 241 AGCAGATACCAAGCTAAAGGGCGATCTCAACAAAGAGATGTCTCTGATCATMAAGAT 300  
QY 340 GTGACTCTGATGATCCATGAGGACTGTGCTGAGATACAGTTCCCTGGTCTTATGAT 399  
Db 301 GTGACTCTGATGATCCATGAGGACTGTGCTGAGATACAGTTCCCTGGTCTTATGAT 360  
QY 400 GATTAATAATTAAGATGTAATTAAGATCAATCAAGACAGCCAGCTCACTCCAGCTCAGACT 459  
Db 361 GATTAATAATTAAGATGTAATTAAGATCAATCAAGACAGCCAGCTCACTCCAGCTCAGACT 420  
QY 460 GCCCATGAGGACTCTTACTACAGCTTCTCCAAAGACCTTAAACAGAGAGAAATGATTC 519  
Db 421 GCCCATGAGGACTCTTACTACAGCTTCTCCAAAGACCTTAAACAGAGAGAAATGATTC 480  
QY 520 GAGACACAGACACTGTGAGACCTCCCATTAATAAGAAAGAAATTTCCATGAGGCT 579  
Db 481 GAGACACAGACACTGTGAGACCTCCCATTAATAAGAAAGAAATTTCCATGAGGCT 540  
QY 580 GATGAATAATTAAGATGTAATTAAGATCAATCAAGACAGCCAGCTCACTCCAGCTCAGACT 639  
Db 541 GATGAATAATTAAGATGTAATTAAGATCAATCAAGACAGCCAGCTCACTCCAGCTCAGACT 600  
QY 640 TCTGCTGGGTTGACCTGTGCACTTATCATTTGTTCTTAAATGATTCCTGT 699  
Db 601 TCTGCTGGGTTGACCTGTGCACTTATCATTTGTTCTTAAATGATTCCTGT 660  
QY 700 AAGAAATAAGATTAATGATGATGAGCTTATCACTGAGCAATGCTCCAGAGAGG 759  
Db 661 AAGAAATAAGATTAATGATGATGAGCTTATCACTGAGCAATGCTCCAGAGAGG 720  
QY 760 TTGGCAATGACAGACAGTCAAGATTCGCTGAGGAAATATCTACACATGAGAG 819  
Db 721 TTGGCAATGACAGACAGTCAAGATTCGCTGAGGAAATATCTACACATGAGAG 780  
QY 820 AACGTATATGATGAGAAATTCATTAATGATGATGATGATGATGATGATGATGATGAT 879  
Db 781 AACGTATATGATGAGAAATTCATTAATGATGATGATGATGATGATGATGATGATGAT 840  
QY 880 TCC 882  
Db 841 TCC 843

## RESULT 7

US-10-252-131-23  
Sequence 0, Application US/10252131  
Publication No. US20030158399A1  
GENERAL INFORMATION:  
APPLICANT: Levins, Douglas A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/252,131  
FILING DATE: 20-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,583  
FILING DATE: 01-MAR-1996  
APPLICATION NUMBER: US 08/487,748  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/398,633  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: Human 200 gene nucleotide  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2236 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 42...944  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-10-252-131-23  
Query Match 18.4%; Score 499.8; DB 12; Length 2236;  
Best Local Similarity 68.5%; Pred. No. 2.4e-140;  
Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps 9;  
QY 17 CGAATTCCTCCCAAGTACTGATGTTTTCAGGTTTACCTCACTGTGCTGCTGC 76  
Db 19 CTGACTTTCTTCTGCAAGCTCCATGTTTTCATCTTCCCTTGTGCTGTCTGCTGC 78  
QY 77 TCTGCAATCTACTTCTGCAAGTCTTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 136  
Db 79 TCTGCTGTCTACTTACTTCAAGTCTTCAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGT 138  
QY 137 ATGCTATCTGCTGCAAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 196  
Db 139 ATGCTATCTGCTGCTGCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 198  
QY 197 GGGGCAAGGATCTGCTCTTGTGTCACAGTGTACCAATGATGATGATGATGATGATGATGAT 256  
Db 199 GGGGCAAGGATCTGCTCTTGTGTCACAGTGTGCAACGTTGCTCAGAGATGATGAA 258  
QY 257 GAAATGTGACATATCAGAAATCCAGAGTACAGCTTAAGGCGATCCAAAGAGAG 316  
Db 259 GGAATGTGATTAATGACATC--CAGTACTGTGTAATGAGGATTTCCCAAGAGAG 315  
QY 317 ATGCTCTGTGATCAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 376  
Db 316 ATGCTCTGTGATCAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375  
QY 377 TACAGTTCCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436  
Db 376 TCCAAATCCAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435

437 CCAAGTCTACCTCCAGCTAGACTGCCCATGGGAGCTTACTACAGCTTCTCCAGAACCC 496  
 436 CCAAGTCTACCTCCAGCTAGACTGCCCATGGGAGCTTACTACAGCTTCTCCAGAACCC 495  
 497 TTAACGAGGAGGAATGG---TTAGAGACACACACCTGGTACCTCTCCATATAACA 553  
 496 TTACACACGAGGAGATGGGCCCGACAGACACACACTGGGAGCTCTCCCTGATATAA 555  
 554 ATGGAACAAATTTCCAC-----ATGGCTGATGAA 586  
 556 ATCTAACACAAATTTCCACATTTGGCCATAGAGTACGGAGCTTGAATTTGGCCATGACT 615  
 587 TTAAGAGCTCTGGAGAAACGATCAGACTGCTATCCATTTAGAGTGGAGTCTGTGCTG 646  
 616 TACGGGACTCTGGAGCAACATCAGAAATAGGCATCTACATCGGACAGGAGATCTGTGCTG 675  
 647 GGTTCACCTTGGCACTTATCATGTGTCTTAACTTAAATGTAATTCCTGTAAAGAAA 706  
 676 GGGCTGCTCTGCTCTTATCTTCCGGCGCTTAAATTTTCAAAATGTAATTCCTATACCAAG 735  
 707 AGAAGTTATGAGTTGAGCTTATTAACCTGAGCACTGGCCAGAGGAGGTTGGCA 766  
 736 AGAAGTATACAGAAATTTAAGCTCATCTTTGGCCAACTCCCTCCAGGATTTGGCA 795  
 767 ATGAGAGCAGAGTCAAGATTGCTCTGAGAAATATCTACACATCGAGAGAAAGTAT 826  
 796 ATGAGTATGAGAGGAAATTTGGCTCAGAAAGAAATCTATATCTATGAGAGAAAGTAT 855  
 827 ATGAGTATGAGAGAAATTTCAATGAGTCTACTGCTACAGCAAGCAAGCCTATGAC 886  
 856 ATGAGTATGAGAGAGCCCAATGATTTATTTGCTATGCTAGCAGCAGGAGCAACCTCTAC 915  
 887 GGCCTCTGAGCTGACCTTTTAAAGG-----TGGCTTCAATTTTGAATTTGG 935  
 916 AACCTTTGGTGTGCTTTGCAATGCTATGATCAACACCTTATTTTGAATTTGGT 975  
 936 TATTTCCCTTTTGGAAAAATATGATATGATGATGATGATGATGATGATGATGATGAT 995  
 976 GTTTGTCTTTTTCAGAAATCTATGATGATGATGATGATGATGATGATGATGATGAT 1033  
 996 CCAAGGCTACGAGAAAGAGTTCAGATTTTCTGGGAGTAAATTAATCTACAGAGGAGTTC 1055  
 1034 CCAAGGCTACGAGAGAGTTCAGATTTTCTGGGAGTAAATTAATCTACAGAGGAGTTC 1092  
 1056 GACTCTAATCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112  
 1093 AACCTGGAGCTGCACTGAACTTAAACAGGATGATGATGATGATGATGATGATGATGAT 1152  
 1113 GATCTCCAGCTCCAGAGACT--TCAATCATGCTGTTTGAAGTCACTGCTGCTTT--CATA 1169  
 1153 AGTACCAACCCAGAGACTGTTATCATGATGATGATGATGATGATGATGATGATGATGAT 1212  
 1170 CATTAGGAATGTTAGTGTGATGCTTTG 1198  
 1213 CACTAGGAATTTCTGACGTGGGCTCTG 1241

RESULT 8  
 US-10-004-633-23  
 ; Sequence 23, Application US/10004633  
 ; Publication No. US20030069196A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; APPLICANT: Lloyd, Clare M.  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
 ; FILE REFERENCE: 7853-125  
 ; CURRENT APPLICATION NUMBER: US/10/004,633  
 ; PRIOR FILING DATE: 2001-12-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

US-10-004-633-23  
 ; Sequence 23, Application US/10004633  
 ; Publication No. US20030069196A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Levinson, Douglas A.  
 ; APPLICANT: Lloyd, Clare M.  
 ; APPLICANT: McCarthy, Sean A.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
 ; FILE REFERENCE: 7853-125  
 ; CURRENT APPLICATION NUMBER: US/10/004,633  
 ; PRIOR FILING DATE: 2001-12-04  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/324,986  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-02  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/609,583

17 CGAATTCCTCCCAAGTACTCAATGTTTGAAGCTTACCTCAAGTGTCTGCTGC 76  
 19 CTGACTTTCTTCTGCAAGCTTCATGTTTCAATCTCTTGAATGCTGTCTGCTGC 78  
 77 TCTGCAACTACTACTTCTGCAAGTCTTGAAGATGTTTATAGGTTGAGTTGTTAA 136  
 79 TCTGCTCTCTACTTCTTCAAGTCTTGAAGATGTTTATAGGTTGAGTTGTTAA 138  
 137 ATGCTATCTGCTCTGCAAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 196  
 139 ATGCTATCTGCTCTGCAAGTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 198  
 197 GGGGCAAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
 199 GGGGCAAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258  
 257 GAAATGTCATATTCAGAAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAG 316  
 259 GGAATGTCATATTCAGAAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAG 315  
 317 ATGCTCTCTGATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 376  
 316 ATGCTCTCTGATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 375  
 377 TACAGTCTCTGATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 436  
 376 TCAAAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 435  
 437 CCAAGTCTACCTCCAGCTAGACTGCCCATGGGAGCTTACTACAGCTTCTCCAGAACCC 496  
 436 CCAAGTCTACCTCCAGCTAGACTGCCCATGGGAGCTTACTACAGCTTCTCCAGAACCC 495  
 497 TTAACGAGGAGGAATGG---TTAGAGACACACACCTGGTACCTCTCCATATAACA 553  
 496 TTACACACGAGGAGATGGGCCCGACAGACACACACTGGGAGCTCTCCCTGATATAA 555  
 554 ATGGAACAAATTTCCAC-----ATGGCTGATGAA 586  
 556 ATCTAACACAAATTTCCACATTTGGCCATAGAGTACGGAGCTTGAATTTGGCCATGACT 615  
 587 TTAAGAGCTCTGGAGAAACGATCAGACTGCTATCCATTTAGAGTGGAGTCTGTGCTG 646  
 616 TACGGGACTCTGGAGCAACATCAGAAATAGGCATCTACATCGGACAGGAGATCTGTGCTG 675  
 647 GGTTCACCTTGGCACTTATCATGTGTCTTAACTTAAATGTAATTCCTGTAAAGAAA 706  
 676 GGGCTGCTCTGCTCTTATCTTCCGGCGCTTAAATTTTCAAAATGTAATTCCTATACCAAG 735  
 707 AGAAGTTATGAGTTGAGCTTATTAACCTGAGCACTGGCCAGAGGAGTGGCA 766  
 736 AGAAGTATGAGAGAAATTTAAGCTCATCTTTGGCCAACTCCCTCCAGAGATTTGGCA 795  
 767 ATGAGAGCAGAGTCAAGATTGCTCTGAGAAATATCTACACATCGAGAGAAAGTAT 826

Query Match 18.4%; Score 499.8; DB 14; Length 2236;  
 Best Local Similarity 68.5%; Pred. No. 2.4e-140;  
 Matches 842; Conservative 1; Mismatches 333; Indels 53; Gaps

Db 796 ATGCAATGACGAGGGGATTCCTCTCAGAGAGAAAACATCTATACCATTTGAAGAGCTAT 855  
 QY 827 ATGAATGAGAAATTCMAATGAGTACTGCTACGTCAACAGCCAGCCATCTGTAC 886  
 Db 856 ATGAATGAGAGAGCCCAATGAGTATATGCTATGTCAGACAGAGCCAAACCCCTCAC 915  
 QY 887 CGGCTCTGAGCTCCCACTTTTAAAGC-----TGGCTTCACTTTGACCTTTGG 935  
 Db 916 AACCTTTGGGTTCGCTTTCGATGCAATGCAATGATCCAAACCACTTATTTTGAAGCTTGT 975  
 QY 936 TATTTCCCTTTTGTGAAAACATATGATGATGATGCTGCACTGGCACTCATTTGGAGTTCTGA 995  
 Db 976 GTTTTGTCTTTTTCAGAAACATATGAGCTGTGTCACTGACTGCTT--TTGAGAGTCTGT 1033  
 QY 996 CCACAGCCACTGAGAAAAGAGTTCAGTTTCTGGGGATTAATTAACACAGAGGATTC 1055  
 Db 1034 CCACTGCTATGAGAGAGAGTTCCTCCATTTTCAGAAAGATA--TGACTCATGAGGAATTG 1092  
 QY 1056 GACTGTACTCATGCTACATTGAATGCTCCATTATCCCTGAGTTTCAGAG--ATCG 1112  
 Db 1093 AACCTGGACCTGACGATGAACTTAAACAGGATGTCATTCCTGTATTTTAAACCAACAG 1152  
 QY 1113 GATCTCCCACTCCAGAGACT--TCAATCATGCGTGTGAAGCTCAGCTGCTT--CAT 1169  
 Db 1153 AGTTACCAACCAAGAGACTGTATATCATGATGATGAGCTCAACAGGGCTTTATATA 1212  
 QY 1170 CATTAGGAATGTTAGTGTGATGCTTTG 1198  
 Db 1213 CACTAGGAATTTCTGACGTGGGCTCTCTG 1241

RESULT 9  
 US-10-188-012-30  
 ; Sequence 30, Application US/10188012  
 ; Publication No. US20030124114A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McIntire, Jennifer Jones  
 ; APPLICANT: Umetsu, Dale T.  
 ; APPLICANT: Dekruyf, Rosemarie  
 ; APPLICANT: Kuchroo, Vijay  
 ; APPLICANT: Freeman, Gordon J.  
 ; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of  
 ; FILE OF INVENTION: Use Thereof  
 ; FILE REFERENCE: STAN-235  
 ; CURRENT APPLICATION NUMBER: US/10/188,012  
 ; CURRENT FILING DATE: 2002-07-01  
 ; PRIOR APPLICATION NUMBER: 60/302,344  
 ; PRIOR FILING DATE: 2001-06-29  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 30  
 ; LENGTH: 1116  
 ; TYPE: DNA  
 ; ORGANISM: H. sapiens  
 US-10-188-012-30

Query Match 17.9%; Score 485.8; DB 14; Length 1116;  
 Best Local Similarity 70.0%; Pred. No. 2,7e-136;  
 Matches 759; Conservative 1; Mismatches 278; Indels 47; Gaps 6;

QY 17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGTCTTACCTTCATGTGTCTGTGC 76  
 Db 35 CTGACTTTTCTTTCGAGGCTCAATGTTTTCACATCTTCCCTTGAAGTGTCTGTGC 94  
 QY 77 TGGTGAATCTACTTCTGCAAGGTCAATGGAATGTTTAAAGTTGAGTTGTTAA 136  
 Db 95 TGGTGTCTGCTACTTCTTCAAGGTCTTCAAGGTGGAATATACAGGCGGAGTGGTGA 154  
 QY 137 ATGCTTATCTGCTGCAAGTTTCACTTACCTTCACTGAGGCACTTGTCTTATGTCT 196  
 Db 155 ATGCTTATCTGCTGCTTCTTACACCCCAAGCCGCGGAGGAACCTGTGCGCGTCTGT 214

QY 197 GGGGCAAGGAGATTTCTCTCTTGTGTCACAGTATACCAATGATGTTGTCAGAACTGATGA 256  
 Db 215 GGGGCAAGGAGACCTCTCTGTTTGAATGGAACCTGTGCTCAGAGCTGATGA 274  
 QY 257 GAAATGTGACATATCAGAAATCCAGAGATATACAGCTTAAAGGCGATCTCAAGAGAG 316  
 Db 275 GGGATGTGAATTTATGGAATC---CAGATATCTGCTTAAATGGGATTTCCGCAAGAG 331  
 QY 317 ATGTCTCTGATCATTAAGAAATGAGTCTGTGATGACATAGGACCTTACTGTGAGGA 376  
 Db 332 ATGTGCTCTGACATATGAGAAATGATGATCTGACAGACAGTGGATCTTCTGTGCGGA 391  
 QY 377 TACAGTTCCCTGTCTTATGAATGAATTAATACTGAATATGACATTAAGACGAG 436  
 Db 392 TCCAAATCCAGCAATTAATGAATGAATTAATTAACCTGAAGTTGTCATTAACGAG 451  
 QY 437 CCAAGGTCACTCAGCTCAGACTGCCATAGGGGACTCTACTACGCTTCCCAAGAACCC 496  
 Db 452 CCAAGGTCACTCAGCTCAGACTGCCATAGGGGACTCTACTACGCTTCCCAAGAACCC 511  
 QY 497 TAAACAGGAGAAATGAG--TTCAGAGACACAGACTGTGACCTTCATATATACA 553  
 Db 512 TTACACAGAGGAGACATGAGCCAGACAGACACAGACATGGGGAGCTTCCGATATATA 571  
 QY 554 ATGGAACAAAATTTCCAC-----ATGGCTGATGA 586  
 Db 572 ATCTAACCAAAATTCACATTTGGCCAAATGAGTTACGGGACTCTTATGTTGGCCAAAGACT 631  
 QY 587 TTAAGGACTCTGAGAGAAAGATCAGAACTGTATCCAAATGAGAGGAGTCTGTCTG 646  
 Db 632 TACGGGACTCTGAGAGAACATCAGAAATGAGGATCTACATCGAGAGGAGATCTGTGCTG 691  
 QY 647 GGTGAACCTGGACTTATCATGTTGTGTCTTAACTTAAATGATATCTCTGTAAGAAA 706  
 Db 692 GGCTGCTCTGCTCTTATCTTCGGGCTTTAATTTCAATGATATCTCATATAGCAAG 751  
 QY 707 AGAAGTTATGAGTTTGAAGCTTATTAACCTGGCCAACTTGCCTCAGAGGAGTTGCA 766  
 Db 752 AGAAGTTATGAGTTTGAAGCTTATTAAGCTTCTTTGGCCAACTTGCCTCAGAGTTGCA 811  
 QY 767 ATGCAAGACAGTCAAGATTCGCTGAGAGAAATATCTTACACCATGAGAGAAAGTAT 826  
 Db 812 ATGCAATACAGAGGAGAAATTCGCTCAGAGAAAGAAACATATACATTTGAAGAAAGTAT 871  
 QY 827 ATGAATGAGAAATTCAAATGAGTACTGCTGCTGCTCAACAGCCAGCCATCTGTAC 886  
 Db 872 ATGAATGAGAGAGCCCAATGAGTATTTATGCTATGTCAGACAGGAGCAACCTCTCAC 931  
 QY 887 CGGCTCTGAGCTGCCCTTTTAAAGC-----TGGCTTCACTTTGACCTTTGG 935  
 Db 932 AACCTTTGGGTTCGCTTTCGATGCAATGATCCAAACCACTTATTTTGAAGCTTGT 999  
 QY 936 TATTTCCCTTTTGTGAAAACATATGATATGTCATCTTGGCAACCTGATTTGAGTTCTGA 995  
 Db 992 GTTTTGTCTTTTTCAGAAACATATGAGCTGTGTCACTGACTGCTT--TTGAGGTTCTGT 1049  
 QY 996 CCACAGCCACTGAGAAAAGAGTTCAGTTTCTGGGGATTAATTAACACAGAGGATTC 1055  
 Db 1050 CCACTGCTATGAGAGAGAGTTTCCATTTTCAGAAAGATA--TGACTCATGAGGAATTG 1108  
 QY 1056 GACTG 1060  
 Db 1109 AACTG 1113

RESULT 10  
 US-10-188-012-32  
 ; Sequence 32, Application US/10188012  
 ; Publication No. US20030124114A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McIntire, Jennifer Jones  
 ; APPLICANT: Umetsu, Dale T.  
 ; APPLICANT: Dekruyf, Rosemarie



```

; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-188-012-32

```

```

Query Match      17.9%; Score 485.8; DB 14; Length 1116;
Best Local Similarity 70.0%; Pred. No. 2.7e-136;
Matches 759; Conservative 1; Mismatches 278; Indels 47; Gaps 6;

```

```

QY 17 CGGATTTCCCTCCCAAGTACTGTTTCAAGTCTTACCTCAACTGTGCTGCTGC 76
DB 35 CTGACTTTCTTCTGCAAGCTTCATGTTTCAATCTTCTTGACTGTGCTGCTGC 94
QY 77 TCTGCAACTACTACTCTTGAAGATGTTAAGATGTTAAGTGTGTTGTTGTTAA 136
DB 95 TCTGCTGCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 154
QY 137 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196
DB 155 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
QY 197 GGGGGAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
DB 215 GGGGGAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274
QY 257 GAATGTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 316
DB 275 GGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
QY 317 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376
DB 332 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 377 TACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB 392 TCCAAATCCGAGCATATGATGATGATGATGATGATGATGATGATGATGATG 451
QY 437 CCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB 452 CCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
QY 497 TAACACGAGAGAAATG---TTGAGAGACAGACACTGCTGCTGCTGCTGCTGCT 553
DB 512 TTACACACGAGAGAAATG---TTGAGAGACAGACACTGCTGCTGCTGCTGCTGCT 571
QY 554 ATGGAACAAATTTTCCAC-----ATGGCTGATGAA 586
DB 572 ATCTAACACAAATTTTCCAC-----ATGGCTGATGAA 631
QY 587 TTAAGGACTCTGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 646
DB 632 TACGGAGCTCTGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 691
QY 647 GGTGACCTGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 706
DB 692 GGTGACCTGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 751
QY 707 AAGAGTATGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 766
DB 752 AAGAGTATGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 811

```

```

QY 767 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
DB 812 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871
QY 827 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 886
DB 872 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931
QY 887 GCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 935
DB 932 AACCTTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
QY 936 TATTTCCCTTTGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 995
DB 992 GTTTGTTGCTTTTGAAGAAATGATGATGATGATGATGATGATGATGATGAT 1049
QY 996 CCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1055
DB 1050 CCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1109
QY 1056 GACTG 1060
DB 1109 AACTG 1113

```

```

RESULT 11
US-10-252-131-37
; Sequence 37, Application US/10252131
; Publication No. US20030158399A1
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/252,131
; FILING DATE: 20-SEP-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cortuzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 37:

```





Db 598 AGAATAGCATCTACATCGAGCAGGAGATCTGTGCGGCTGCTGCTCTTAATCTTC 657  
Qy 670 GGTGCTTAATCTTAATGATGATTCCTGTAAGAAAAAGATTATCGAGTTTGACCTT 729  
Db 658 GGGCTTTAATTTTCAATGATGATTTCTATAGCAAGAAAGATACAGATTTAAAGCTTC 717  
Qy 730 ATTACACTGGCCAACTTGGCTTCAGAGAGGTTGGCAATATGACAGACAGTCAAGATTGCG 789  
Db 748 ATCTCTTTGGCAACCTCCCTCCCTCAGGATTGGCAATATGCAATGAGGAAATTCGC 777  
Qy 790 TCTGAGAAATAATATCTACACCATCGAGGAAACGTTATGAAGTGAAGATTTCAATGAG 849  
Db 778 TAGAAGAAACATCTATACATTCATTAAGAAAGATATGAAGTGAAGGAGCCCAATGAG 837  
Qy 850 TACTCTGCTAGCTCAACAGCAGCAGCCATCTGACCGCTCTGAGCTGCCACTTTTA 909  
Db 838 TATTATGCTATGTCAGCAGCAGGAGCAACCTTCAACCTTTGGGTTGTGCTTTGCA 897  
Qy 910 AGGC 913  
Db 898 ATGC 901

RESULT 13  
US-09-529-063-36  
; Sequence 36, Application US/09529063  
; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; CURRENT FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514  
; PRIOR FILING DATE: 1998-10-06  
; PRIOR APPLICATION NUMBER: JP 9-274674  
; PRIOR FILING DATE: 1997-10-07  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 2091  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (53)..(115)  
; NAME/KEY: mat\_peptide  
; LOCATION: (116)..(607)  
; NAME/KEY: CDS  
; LOCATION: (53)..(607)  
US-09-529-063-36

Query Match 12.1%; Score 328.8; DB 10; Length 2091;  
Best Local Similarity 62.4%; Pred. No. 2.2e-88;  
Matches 750; Conservative 1; Mismatches 298; Indels 153; Gaps 9;

Qy 17 CGGATTTCCCTCCCAAGTACTATGTTTTCAGGTCTTACCTCAACGTGTCTGCTGC 76  
Db 30 CTGACTTTTCTCTGCAAGCTCATGTTTTCACATCTTCCCTTGACTGTGCTGCTGC 89  
Qy 77 TGTGCAACTACTACTTTCGAGGTCTATTGGAAGATGTTTAAAGTTGAGTTGTTAA 136  
Db 90 TGTGTGTACTACTTACAAAGTCTCAGAACTGAAATACAGAGGAGGTGCTGACAA 149  
Qy 137 ATGCTATCTGCGCTGAGTACACTCTTACTTACATCTGAGCACTGTGCTATGCT 196  
Db 150 ATGCTATCTGCGCTGAGTACACTCTTACTTACATCTGAGCACTGTGCTATGCT 209  
Qy 197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGATGCTCAGAACTGATGAA 256

Db 210 GGGGCAAGGAGCTGTCTGTCTGTTTGAATGTGGCAAGTGTGCTCAGACTGATGAA 269  
Qy 257 GAATGTGACATATGCAAAATCCAGCAGATACAGCTAAAGGGGAGTCCAAAGAGG 316  
Db 270 GGGATGTAAATTTATGACATC---CAGATCTGCTAAATGGGATTTCCGCAAGAGG 326  
Qy 317 ATGTCTCTGTATCTAAAGAAATGTACTGTGATGACATAGGAGCTTACTGTCTCAGGA 376  
Db 327 ATGTCTCTGTATCTAAAGAAATGTACTGTGATGACATAGGAGCTTACTGTCTCAGGA 386  
Qy 377 TACAGTTCCCTGCTTTTATGAATGATTAATAAATTAAGAACTGAATTAAGACAG 436  
Db 387 TCCAAATCCCAAGGCAATATGATGATGAAATTTAACTGAATGTGTGATCAATCAACAG 446  
Qy 437 CCAAGTCACTCAGCTCAGACTGCGCCATGGGGACTTACTACAGCTTCCCAAGACCC 496  
Db 447 CCAAGTCACTCAGCTCAGACTGCGCCATGGGGACTTACTACAGCTTCCCAAGAGTC 506  
Qy 497 TAACCAAGGAGAGAAATG---TTCAGAGACACAGACACTGTGACCTCCATTAATCA 553  
Db 507 TTACCAAGGAGGAGCATGGCCACAGACAGACAGACACTGGGAGCTCCCTGATATTA 566  
Qy 554 ATGAAACAAAATTTCCACATGGGCTGATGAATTAAGGACTTGGAGAAACGATCAGA 613  
Db 567 ATCTAACAA-----  
Qy 614 CTGCTATCACATTTGAGATGGGAGTCTGTGCTGGTTGACCTTGCACTTATCATTTGTC 673  
Db 575 -----  
Qy 674 TCTTAATCTTAATGATGATTCCTGTAAGAAAAAGATTATCGATTGAGCTTATTA 733  
Db 575 -----  
Qy 734 CACTGGCAACTTGTCTCCAGAGGTTGGCAATATGACAGACAGTCAAGATTCCCTG 793  
Db 620 CTTTGGCCAACTCCCTCCCTCAGAGATGGCAAAATGACAGTGAAGGAAATTCCTCAG 679  
Qy 734 AGGAAATATCTACACCATTCAGAGAGACGTATGAAGTGAAGTGAATTAATGATGAT 853  
Db 680 AAGAAACATCTACACCATTCAGAGAGACGTATGAAGTGAAGTGAAGTGAAGTGAAT 739  
Qy 854 ACTGCTAGTCAACAGCAGCAGCCATCTGACCCCTCTGACCTGCACTTTTAAAGGC 913  
Db 740 ATTGCTATGATGACAGCAGGAGCAACCTCACAACCTTTGGGTGTGCTTGAATGC 799  
Qy 914 -----TGGCTTCAATTTCTGACTTGTGATTTCCCTTTTGAAAACTATGTGA 962  
Db 800 CATAGATCCAAACCACTTATTTTGAAGCTTGGTGTGTTGCTTTTCAAGAACTATGAGC 859  
Qy 963 TATGTCACTTGGCAACCTCATTTGAGGTTCGACCAAGCCACTGAGAAAGATTCAG 1022  
Db 860 TGTGTACCTGACGTGTT--TTGAGAGTTCGTCTCAGCTATGAGAGCAAGATTTTCCA 917  
Qy 1023 TTTTCTGGGATTAATTAATCACTCAAGGAGATTCAGCTGATTACTCATGCTACATTTGAATG 1082  
Db 918 TTTTCAAGAAATTA--TGACTCACTGAGAAATTAAGTGGAGCTGCACTGAATTAACA 976  
Qy 1083 CTCATTTTATCCCTGAGTTTCAAGG--ATGCAATCTCCCACTCAGAGACT--TCATC 1138  
Db 977 GGCATGATGATGCTCTGTATTTAAGCAACAGATTTCCCAACCAAGAGATGTTAAT 1036  
Qy 1139 ATGCGTGTGAAGTCACTGCTGCTTC--ATACATTAAGAAATGATGATGATGATGAT 1196  
Db 1037 ATGCAATTTAGAGTCAACAGGCTTTTATATATACCTGAGAAATTTCTGACGTGGGTCTC 1096  
Qy 1197 TG 1198  
Db 1097 TG 1098

RESULT 14  
US-10-414-378-36

```
/ Sequence 36, Application US/10414378
/ Publication No. US20030165981A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ FILE OF INVENTION: THE BOTH
/ FILE REFERENCE: 058769
/ CURRENT APPLICATION NUMBER: US/10/414,378
/ CURRENT FILING DATE: 2003-04-16
/ PRIOR APPLICATION NUMBER: US/09/529,063
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 36
/ LENGTH: 2091
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: (53)..(115)
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: (116)..(607)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (53)..(607)
US-10-414-378-36
```

```
Query Match 12.1%; Score 328.8; DB 12; Length 2091;
Best Local Similarity 62.4%; Pred. No. 2.2e-88;
Matches 750; Conservative 1; Mismatches 298; Indels 153; Gaps 9;
```

```
QY 17 CGGATTTCCCTCCCAAGTACTCATGTTTTCAGGCTTTACCTCACTGTCCTGCTGC 76
DB 30 CTGACTTTTCTTCGCAAGCTCCATGTTTTCACATCTTCCCTTGACTGTCCTGC 89
QY 77 TGCCTCACTACTACTTGAAGGTCATTTGGAAGTGTATTAAGTTGAGTTGTTGTTAA 136
DB 90 TGTCTGCTACTACTTACAAAGTCTCAAGTGTGAATACAGAGCGAGGTCGTTACA 149
QY 137 ATGCTATCTGCCCTGCAAGTTCACCTTACTCATCTGGGACACTTGTGCTATGTCT 196
DB 150 ATGCTATCTGCCCTGCTTCTTACACCCAGCCGCCAGGGAACCTGTGCCGTGCT 209
QY 197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTACCAATGAGTTGCTCAGAACTGATGAA 256
DB 210 GGGGCAAGGATTTCTGCTTGTGTCACAGTGTGCAACGTTGCTCAGGACTGATGAAA 269
QY 257 GAAATGTACATATCAGAAATCCAGAGATACAGCTAAAGGGGATCTCAACAAAGAG 316
DB 270 GGAATGTATATTTGAGATC--CAGATATGCTAAATGGGATTTCCGCAAGGAG 326
QY 317 ATGTGTCTGTATCAATTAAGATGATCTGTGATGACATGAGGACCTTACTGTGACGA 376
DB 327 ATGTGTCTGTATCAATTAAGATGATCTGTGATGACATGAGGACCTTACTGTGACGA 386
QY 377 TACAGTTCCCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436
DB 387 TCCAAATCCAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
QY 437 CCAAGGCTCTCAGCTCAGCTGCGCCAGGAGCTTACTACAGCTTCTTCCAAAGACC 496
DB 447 CCAAGGCTCTCAGCTCAGCTGCGCCAGGAGCTTACTACAGCTTCTTCCAAAGATGC 506
QY 497 TAAACAAGAGAGAAATG--TTACAGAGACACAGACACTGTGACCTTCATATATACA 553
DB 507 TTACCAACAGGAGATGCGCCAGCAGACAGACAGACACTGTGAGGAGCTTCCCTGATATA 566
```

```
QY 554 ATGGAACAAAATTTCCACATGGGCTGATGAATTAAGACTGTGGAGAAACATGACGA 613
DB 567 ATCTCA----- 574
QY 614 CTGCTATCCACATTTGAGTGGAGTCTGTGCTGGTTGACCTTGACATTATCATTTG 673
DB 575 ----- 574
QY 674 TCTTAATCTTAATGATTTCTGTGAGAAAAAGATTATGAGCTTTGAGCTTATTA 733
DB 575 -----GTAATTTCTATGACAAAGAGAAATGATGAAATTAAGCTCATCT 619
QY 734 CACTGGCCAACTTGCTCCAGAGAGGTTGGCAATTCAGAGACAGTCTGCTG 793
DB 620 CTTTGGCCAACTTGCTCCAGAGAGGTTGGCAATTCAGAGACAGTCTGCTG 679
QY 794 AGGAAATATCTACACATCGAGAGAGATATGATGAGGAAATTCAAATGATGACT 853
DB 680 AAGAAACATCTATACCATTTGAGAGAGATATGATGAGGAAATTCAAATGATGACT 739
QY 854 ACTGCTACGTCACACAGCAGACCATCTGACCCGCTCTGACCTTGAACCTTTAAAGC 913
DB 740 ATTGCTATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 799
QY 914 -----TGCTTCAATTTCTGACTTTGATTTTCTTCTTCTTCTTCTTCTTCTTCT 962
DB 800 CATGATCCACACCTTATTTTGAAGCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCT 859
QY 963 TATGCTACTTGGCACTCATGATGAGGTTCTGACACAGCAGCAGCAGCAGCAGCAGCAG 1022
DB 860 TGTGTACCTGACTGTT--TTGAGAGTTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 917
QY 1023 TTTTCTGGGATTAATTAATCAACAGGAGTTGACTGATCACTGATCACTGATGAAATG 1082
DB 918 TTTTCAAGAGATTA--TGACTCAGATGGAATGGAATGGAATGGAATGGAATGGAATG 976
QY 1083 CTCATTTTATCCCTGATGTTGAGG--ATGGAATCTCCCACTCCAGAGACT--TCAATC 1138
DB 977 GGGATGTCAATGCTCTGTATTTAAACCAACAGAGTTACCAACCAAGAGACTGTTAAATC 1036
QY 1139 ATGCTGTGGAAGCTCACTGTCCTTTC--ATCATTTAGAAATGTTAGTGTGATCTT 1196
DB 1037 ATGATGTTTGAAGCTCAACAGGCTTTTATATATACATAGGAATTTCTTGAAGGCTTC 1096
QY 1197 TG 1198
DB 1097 TG 1098

RESULT 15
US-09-529-063-35
/ Sequence 35, Application US/09529063
/ Patent No. US20020102542A1
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSHIMA, DAICHI
/ APPLICANT: SHIBAYAMA, SHIRO
/ APPLICANT: TADA, HIDEAKI
/ TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
/ FILE OF INVENTION: THE BOTH
/ FILE REFERENCE: 058769
/ CURRENT APPLICATION NUMBER: US/09/529,063
/ CURRENT FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: PCT/JP98/04514
/ PRIOR FILING DATE: 1998-10-06
/ PRIOR APPLICATION NUMBER: JP 9-274674
/ PRIOR FILING DATE: 1997-10-07
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 35
/ LENGTH: 555
/ TYPE: DNA
/ ORGANISM: Homo sapiens
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 23:34:40 ; Search time 290.42 Seconds  
(without alignments)  
10161.809 Million cell updates/sec

Title: US-10-004-633-37

Perfect score: 903  
Sequence: 1 atgtttccatcttccctt.....gtctgcttgcattgcca 903

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 216961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	903	100.0	903	US-10-252-131-37	Sequence 37, Appl
2	903	100.0	903	US-10-004-633-37	Sequence 37, Appl
3	903	100.0	1116	US-10-188-012-10	Sequence 30, Appl
4	903	100.0	2236	US-10-252-131-23	Sequence 0, Appl
5	903	100.0	2236	US-10-004-633-23	Sequence 23, Appl
6	901.4	99.8	1116	US-10-188-012-32	Sequence 32, Appl
7	585	64.8	2091	US-09-529-063-36	Sequence 36, Appl
8	585	64.8	2091	US-10-414-378-36	Sequence 36, Appl
9	522.8	57.9	555	US-09-529-063-35	Sequence 35, Appl
10	522.8	57.9	555	US-10-414-378-35	Sequence 35, Appl
11	456	50.5	2710	US-10-252-131-8	Sequence 8, Appl
12	456	50.5	2710	US-10-004-633-8	Sequence 8, Appl
13	456	50.5	2725	US-10-188-012-10	Sequence 10, Appl
14	455	50.4	529	US-10-040-739-415	Sequence 415, Appl
15	442.6	49.0	843	US-10-252-131-22	Sequence 22, Appl
16	442.6	49.0	843	US-10-004-633-22	Sequence 22, Appl

17	440.8	48.8	862	US-10-188-012-12	Sequence 12, Appl
18	396.6	43.9	1203	US-09-966-546-21	Sequence 21, Appl
19	396.6	43.9	1203	US-09-966-545-21	Sequence 21, Appl
20	396.6	43.9	1203	US-09-965-212-21	Sequence 21, Appl
21	396.6	43.9	1203	US-10-189-940-21	Sequence 21, Appl
22	333.6	36.9	375	US-10-189-940-144	Sequence 144, Appl
23	82.8	9.2	582	US-10-027-632-258124	Sequence 258124, Appl
24	82.8	9.2	582	US-10-027-632-258124	Sequence 258124, Appl
25	81.6	9.0	582	US-10-027-632-258125	Sequence 258125, Appl
26	81.6	9.0	582	US-10-027-632-258126	Sequence 258126, Appl
27	81.6	9.0	582	US-10-027-632-258125	Sequence 258125, Appl
28	81.6	9.0	582	US-10-027-632-258126	Sequence 258126, Appl
29	72.6	8.0	1079	US-10-188-012-24	Sequence 24, Appl
30	72.6	8.0	1080	US-10-188-012-20	Sequence 20, Appl
31	72.6	8.0	1080	US-10-188-012-20	Sequence 20, Appl
32	72.6	8.0	1095	US-10-188-012-26	Sequence 26, Appl
33	72.6	8.0	1098	US-10-188-012-22	Sequence 22, Appl
34	72.6	8.0	1099	US-10-188-012-28	Sequence 28, Appl
35	66.4	7.4	918	US-10-188-012-2	Sequence 2, Appl
36	66.4	7.4	1032	US-10-188-012-14	Sequence 14, Appl
37	66.4	7.4	1032	US-10-188-012-16	Sequence 16, Appl
38	65.4	7.2	849	US-10-188-012-4	Sequence 4, Appl
39	65.2	7.2	571	US-09-739-907-20	Sequence 20, Appl
40	64.8	7.2	398	US-09-960-352-7732	Sequence 7732, Appl
41	64.8	7.2	418	US-09-960-352-1063	Sequence 1063, Appl
42	64.4	7.1	1156	US-10-188-012-34	Sequence 34, Appl
43	64.4	7.1	1156	US-10-188-012-36	Sequence 36, Appl
44	64.4	7.1	1282	US-09-813-153-68	Sequence 68, Appl
45	64.4	7.1	1380	US-10-037-270-59	Sequence 59, Appl

#### ALIGNMENTS

RESULT 1  
US-10-252-131-37  
Sequence 37, Application US/10252131

Publication No. US20030158399A1

GENERAL INFORMATION:

APPLICANT: Levinson, Douglas A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/252,131

FILING DATE: 20-Sep-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/609,583

FILING DATE: 01-MAR-1996

APPLICATION NUMBER: US 08/487,748

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/398,633

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864





```

Db 481 GAGACACAGACCTGGGAGCCTCCCTGATATTAATCAACAATATCCACATTTGGCC 540
Qy 541 AATAGATTACGGGACTCTAGATTGGCCAAATGCTTAGGGGACTCTGGAGCAACATCGA 600
Db 541 AATAGATTACGGGACTCTAGATTGGCCAAATGCTTAGGGGACTCTGGAGCAACATCGA 600
Qy 601 ATAGGCACTTACATCGAGAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 660
Db 601 ATAGGCACTTACATCGAGAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 660
Qy 661 GCTTTAATTTCAATGATGATTTCTCATACGAAAGAAATACAGATTTAAAGCTCATC 720
Db 661 GCTTTAATTTCAATGATGATTTCTCATACGAAAGAAATACAGATTTAAAGCTCATC 720
Qy 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGAGATGAGCAAGGAAATTCCTCA 780
Db 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGAGATGAGCAAGGAAATTCCTCA 780
Qy 781 GAAGAAAACATCTATACCATTTGAAGAAAGATATATGAGAGAGAGCCCAATGATAT 840
Db 781 GAAGAAAACATCTATACCATTTGAAGAAAGATATATGAGAGAGAGCCCAATGATAT 840
Qy 841 TATTGCTATGTCAGAGAGGAGCAACCTTCAACCTTTGGGTTGTGCTTTGCAATG 900
Db 841 TATTGCTATGTCAGAGAGGAGCAACCTTCAACCTTTGGGTTGTGCTTTGCAATG 900
Qy 901 CCA 903
Db 901 CCA 903

```

RESULT 3

```

US-10-188-012-30
; Sequence 30, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McInitive, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-188-012-30

```

```

Query Match 100.0%; Score 903; DB 14; Length 1116;
Best Local Similarity 100.0%; Pred. No. 4,6e-252;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ATGTTTTCATCTCTCCCTTTGAGTGTGCTGCTGCTGCTGCTACTACTTACAAAGG 60
Db 58 ATGTTTTCATCTCTCCCTTTGAGTGTGCTGCTGCTGCTGCTACTACTTACAAAGG 117
Qy 61 TCTTCAGAGTGAATACAGAGCGGAGGTGCTGCTGCTGCTGCTGCTGCTTCTAC 120
Db 118 TCTTCAGAGTGAATACAGAGCGGAGGTGCTGCTGCTGCTGCTGCTGCTTCTAC 177
Qy 121 ACCCAGCGCGCCAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 178 ACCCAGCGCGCCAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 237

```

```

Qy 181 TTGTAATGTGCAACGTGTGCTCAGAGCTGATGAAAAGGATGTGAATTTATGACATCC 240
Db 238 TTGTAATGTGCAACGTGTGCTCAGAGCTGATGAAAAGGATGTGAATTTATGACATCC 297
Qy 241 AGATACCTGCTAAATGGGGAATTTCCGAAAGAGATGTGTCTCTGACCATAGAGATGTG 300
Db 298 AGATACCTGCTAAATGGGGAATTTCCGAAAGAGATGTGTCTCTGACCATAGAGATGTG 357
Qy 301 ACTCTAGAGACAGTGGGATCTACTGCTGCGGATCCAAATCCAGAGATATGATAT 360
Db 358 ACTCTAGAGACAGTGGGATCTACTGCTGCGGATCCAAATCCAGAGATATGATAT 417
Qy 361 GAAATATTTAACTGAATGTTGTCATCAAAACAGCCAGGATGACCCCTGACCCGACTTG 420
Db 418 GAAATATTTAACTGAATGTTGTCATCAAAACAGCCAGGATGACCCCTGACCCGACTTG 477
Qy 421 CAGAGAGCTTCACTGACGCTTTCCAGAGATGCTTACACAGAGGAGCATGGCCACGA 480
Db 478 CAGAGAGCTTCACTGACGCTTTCCAGAGATGCTTACCAAGAGATGCTTACCAAGGAGCATGGCCACGA 537
Qy 481 GAGACAGAGACCTGGGAGGCTCCCTGATATTAATCTTACACAAATTTCCATTTGGCC 540
Db 538 GAGACAGAGACCTGGGAGGCTCCCTGATATTAATCTTACACAAATTTCCATTTGGCC 597
Qy 541 AATAGATTACGGGACTCTAGATTGGCCAAATGCTTACGGGACTCTGAGCAACATCGA 600
Db 598 AATAGATTACGGGACTCTAGATTGGCCAAATGCTTACGGGACTCTGAGCAACATCGA 657
Qy 601 ATAGGCACTTACATCGAGAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 660
Db 658 ATAGGCACTTACATCGAGAGGAGATCTGTGCTGGGCTGGCTCTTATCTTGGCC 717
Qy 661 GCTTTAATTTCAATGATGATTTCTCATACGAAAGAAATATTAAGCTCATC 720
Db 718 GCTTTAATTTCAATGATGATTTCTCATACGAAAGAAATATTAAGCTCATC 777
Qy 721 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGAGATGAGCAAGGAAATTCCTCA 780
Db 778 TCTTTGGCAACCTCCCTCCCTCAGGATTTGGCAATGAGATGAGCAAGGAAATTCCTCA 837
Qy 781 GAAGAAAACATCTATACCATTTGAAGAAAGATATATGAGAGAGAGCCCAATGATAT 840
Db 838 GAAGAAAACATCTATACCATTTGAAGAAAGATATATGAGAGAGAGCCCAATGATAT 897
Qy 841 TATTGCTATGTCAGAGAGGAGCAACCTTCAACCTTTGGGTTGTGCTTTGCAATG 900
Db 898 TATTGCTATGTCAGAGAGGAGCAACCTTCAACCTTTGGGTTGTGCTTTGCAATG 957
Qy 901 CCA 903
Db 958 CCA 960

```

RESULT 4

```

US-10-252-131-23
; Sequence 0, Application US/10252131
; Publication No. US20030158399A1
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```





















**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2003, 12:17:01 / Search time 5933.8 Seconds  
(without alignments)  
11099.996 Million cell updates/sec

Title: US-10-004-633-8  
Perfect score: 2710  
Sequence: 1 ngcgaccacgcgcgcga.....aaaaaaaaagcgccgc 2710

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rpd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	567.2	20.9	633	14	BY741594	BY741594 BY741594
2	552.8	20.4	677	14	BY748679	BY748679 BY748679
3	472.6	17.4	491	10	BE686651	BE686651 uw03c07.x
4	466	17.2	1201	13	BX417090	BX417090 BX417090

5	446	16.5	476	4	BX526231	BX526231 RZPD Mus
6	440	16.2	440	10	BB750705	BB750705 BB750705
7	381	14.1	381	13	BY528967	BY528967 BY528967
8	345.6	12.8	374	9	AT152672	AT152672 ud93e08.r
9	324	12.0	354	9	AT152783	AT152783 ud91b01.r
10	302	11.1	583	13	B0789118	B0789118 1040b02.x
11	292.2	10.8	735	14	CA307836	CA307836 UI-R-FT1-
12	291.8	10.8	863	14	CD520418	CD520418 AGENCOURT
13	289.8	10.7	554	13	B0581947	B0581947 111a10.y
14	286.8	10.6	575	9	AU279844	AU279844 AU279844
15	276	10.2	746	14	CD522575	CD522575 AGENCOURT
16	273.4	10.1	869	14	CD523271	CD523271 AGENCOURT
17	264.8	9.8	350	13	BY185793	BY185793 BY185793
18	247.8	9.1	518	14	CB155731	CB155731 K-EST0214
19	247.6	9.1	395	10	BP386140	BP386140 UI-R-CA1-
20	245	9.0	796	12	B1906710	B1906710 603064430
21	234.6	8.7	572	9	AL709094	AL709094 DKF2P686F
22	232.4	8.6	280	9	AV248049	AV248049 AV248049
23	232	8.6	267	10	BB434450	BB434450 BB434450
24	231.6	8.5	837	10	BG545730	BG545730 602573003
25	213.6	7.9	296	10	BB572107	BB572107 BB572107
26	213.2	7.9	229	9	AV249853	AV249853 AV249853
27	200.2	7.4	519	12	B1834671	B1834671 603090448
28	196	7.2	1370	13	BQ278485	BQ278485 AGENCOURT
29	189.4	7.0	606	9	AV664631	AV664631 AV664631
30	163.8	6.0	535	12	B1541253	B1541253 454905.MA
31	156.8	5.8	168	10	B1706677	B1706677 B1706677
32	150.6	5.6	429	9	AV664630	AV664630 AV664630
33	133.2	4.9	655	13	B0581620	B0581620 111a10.x
34	111	4.1	557	12	BM255779	BM255779 517968.MA
35	95.8	3.5	568	12	BM261288	BM261288 da146d03.
36	95.2	3.5	525	12	B1442322	B1442322 da138e11.
37	95.2	3.5	641	13	B0400478	B0400478 NISC.MP09
38	92.2	3.4	828	14	CD254919	CD254919 AGENCOURT
39	92.2	3.4	634	13	B0238988	B0238988 603321725
40	92.2	3.4	694	9	AJ452225	AJ452225 AJ452225
41	92.2	3.4	776	9	AJ452245	AJ452245 AJ452245
42	92.2	3.4	807	9	AJ445463	AJ445463 AJ445463
43	84.4	3.1	754	14	CB938882	CB938882 IPGCTx13
44	84.2	3.1	754	13	B0320531	B0320531 603850370
45	84.2	3.1	900	13	B0346895	B0346895 603524895

## ALIGNMENTS

RESULT 1  
LOCUS BY741594  
DEFINITION BY741594 RIKEN full-length enriched, adult pancreas islet cells Mus  
musculus CDNA clone C82007B12 5', mRNA sequence.  
ACCESSION BY741594  
VERSION BY741594.1 GI:27166585  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.  
1 (bases 1 to 633)  
Okazaki, Y., Oeato, N., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamana, A., Kiyosawa, H.,  
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogi, A., Schonbach, C.,  
Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,  
Quackenbush, J., Schriml, L., Kanapin, A., Matsuda, H., Batalov, S.,  
Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,  
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,  
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J.,  
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,  
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,  
P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Mikl,  
H., Nagashima, T., Numata, K., Okado, T., Pavani, W.J., Perte, G.,  
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,

Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shima, K., Sultana, K., Takanaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yang, L., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, O., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sato, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, T. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Cells were provided by Hiroo Iwata (Department of Reproductive Materials Field of Tissue Engineering Institute for Frontier Medical Sciences, Sakyo-ku, Kyoto, 606-8507, Japan) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

1. 633

organism="Mus musculus"

molecule="mRNA"

db\_xref="taxon:10090"

clone="CG20007E12"

tissue\_type="pancreas"

cell\_type="islet cells"

dev\_stage="adult"

lab\_host="DH108"

clone\_1b="RIKEN full-length enriched, adult pancreas islet cells"

note="Site 1: Salt, Site 2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken

BASE COUNT 188 a 150 c 139 g 156 t

ORIGIN

Query Match 20.9%; Score 567.2; DB 14; Length 633;  
Best Local Similarity 97.8%; Pred. No. 2.7e-77;  
Matches 575; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

18 GGATTTCCCTCCCAAGTACTGATGTTTTCAGTCTTACCTCACTGTCTGCTGCT 77  
46 GGATTTCCCTCCCAAGTACTGATGTTTTCAGTCTTACCTCACTGTCTGCTGCT 105

78 GCTGCACTACTACTTCTGCAAGTCAATGGAAGTGTATTAAGTTGAGTTGTTAA 137  
106 GCTGCACTACTACTTCTGCAAGTCAATGGAAGTGTATTAAGTTGAGTTGTTAA 165

138 TGCCATATCGCCCTGAGTATACCTCTACCTGAGGACACTTGCTATGCTG 197  
166 TGCCATATCGCCCTGAGTATACCTCTACCTGAGGACACTTGCTATGCTG 225

198 GGGCAAGGATTCCTGCTTGTGTCACAGTACCAATGATGCTCAGAACTGATGA 257  
226 GGGCAAGGATTCCTGCTTGTGTCACAGTACCAATGATGCTCAGAACTGATGA 285

258 AATGTCATATTCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAAGAG 317  
286 AATGTCATATTCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAAGAG 345

318 TGTGTCATATTCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAAGAG 377  
346 TGTGTCATATTCAGAAATCCAGCAGATACCAAGTAAAGGCGATCTCAAGAG 405

378 AAGTTCCTGCTTATGATGATTAATTAATTAAGTAAATTAAGATCAATCAAG 437  
406 AAGTTCCTGCTTATGATGATTAATTAATTAAGTAAATTAAGATCAATCAAG 465

438 CAAGTGCATCCAGCTCAGATCCGATGGGAGCTCTACTACGCTTCTCAAGAAC 497  
466 CAAGTGCATCCAGCTCAGATCCGATGGGAGCTCTACTACGCTTCTCAAGAAC 525

498 AACCAAGGAGAAATGTTGAGAGACAGACACTGGGACCTCCATATAATCAAT 557  
526 AACCAAGGAGAAATGTTGAGAGACAGACACTGGGACCTCCATATAATCAAT 585

558 AACCAAAATTCACATGGGCTGATGAATTAAGACTCTGAGAAAC 605  
586 AACCAAAATTCACATGGGCTGATGAATTAAGACTCTGAGAAAC 633

RESULT 2  
BY748679 677 bp mRNA linear EST 17-DEC-2002  
LOCUS BY748679 RIKEN full-length enriched, NOB-derived CD11c +ve  
DEFINITION dendritic cells Mus musculus cDNA clone F630119G06 5', mRNA  
sequence.  
ACCESSION BY748679  
VERSION BY748679.1 GI:27177605  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGAGAGAGAGGCGCCCAAGTACTGATGTTTTCAGTCTTACCTCACTGTCTGCTGCT prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGATTCGATTAATTAATTAATTCACCTGAGGACACTTGCTATGCTG was cleaved with BamHI and XhoI. Vector: a modified pluscript KS(+) after bulk excision from Lambda PUC 1. Cells were provided by Tomohiro Kono, Department of Animal Science, Tokyo University of Agriculture, Atsugi City, Kanagawa Prefecture, Japan, whose assistance we gratefully acknowledge."



## Tumor Gene Index

## JOURNAL

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:1091408

## FEATURES

High quality sequence stop: 459.

## source

Location/Qualifiers

1.491  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3415596"  
/sex="male"  
/tissue\_type="Spleen"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares mouse 3NBMS"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5].  
TGTACCAATCTGAGGAGGAGCGCGCTGTTTGTGAGGAGGAGCGTCTT  
3); double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through  
three rounds of normalization, and was constructed by  
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 145 a 109 c 94 g 143 t

## ORIGIN

Query Match 17.4%; Score 472.6; DB 10; Length 491;  
Best Local Similarity 99.0%; Pred. No. 9.1e-63;  
Matches 486; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

2199 CAAGGCTCTGCGACATCTATATAGTCAGCAAAATTTCTTGGGAGGACAGTCGTCAC 2258  
491 CAGAGCTCTGCGACATCTATATAGTCAGCAAAATTTCTTGGGAGGACAGTCGTCAC 432  
2259 CAATGATTTCCAGCGGAGGAGCTGAGTTTCATCTGCTTACAGCTGCTGCCAGT 2318  
431 CAATGATTTCCAGCGGAGGAGCTGAGTTTCATCTGCTTACAGCTGCTGCCAGT 372  
2319 GCCCTGATCTGCTGCTGCCATCTATAACAAGATCAAAATTAATAGACCCCGAGTGA 2378  
371 GCCCTGATCTGCTGCTGCCATCTATAACAAGATCAAAATTAATAGACCCCGAGTGA 312  
2379 AAATTAATAGTGAAGGAGGAGTGAAGTCTTGTCAAAATTTTGGATGGGAGGAC 2438  
311 AAATTAATAGTGAAGGAGGAGTGAAGTCTTGTCAAAATTTTGGATGGGAGGAC 252  
2439 TGTGTACATCAGAGCATCTGTTAGTGAAGA-CACAAACCTGTGTACCGTTTTC 2497  
251 TGTGTACATCAGAGCATCTGTTAGTGAAGACCCCAAAACCTGTGTACCGTTTTC 132  
2498 ATGTATGAATTTTGTGTTAGTGTGCTTCTAGCTAGCTGTGAGGAGTCTGCTTCTTA 2557  
191 ATGTATGAATTTTGTGTTAGTGTGCTTCTAGCTAGCTGTGAGGAGTCTGCTTCTTA 132  
2558 GTGTGGATGGAAGGAGGAGCATCTTAACAATTCATTAGAGATTAACAGCTTCTACAG 2617  
131 GTGTGGATGGAAGGAGGAGCATCTTAACAATTCATTAGAGATTAACAGCTTCTACAG 72  
2618 AAGGAGAACTAATCTCAAAATGTTTAAAGTAATTAAGTACTGCGCAAGTACTTGA 2677  
71 AAGGAGAACTAATCTCAAAATGTTTAAAGTAATTAAGTACTGCGCAAGTACTTGA 12  
2678 GCATATAAAAA 2688  
11 GCATATTATA 1

## RESULT 4

## EX417090

## LOCUS

EX417090 1201 bp mRNA linear EST 15-MAY-2003

## DEFINITION

EX417090 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE004YC05

## ACCESSION

EX417090

## VERSION

EX417090.1 GI:30765665

## KEYWORDS

EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1. (bases 1 to 1201)  
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

## AUTHORS

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished  
Contact: Genoscope

## COMMENT

Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 2998.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE004AB03QPI&cluster=2998.r. Contact :  
Feng Liang Email: fliang@life.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE004AB03QPI.  
Location/Qualifiers

## FEATURES

## source

1.1201  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DE004YC05"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## BASE COUNT

312 a 282 c 276 g 295 t 36 others

## ORIGIN

Query Match 17.2%; Score 466; DB 13; Length 1201;  
Best Local Similarity 68.0%; Pred. No. 6.7e-62;  
Matches 723; Conservative 14; Mismatches 280; Indels 47; Gaps 5;

17 CGGATTTCCCTCCCAAGTACATCTTTTTCAGTCTTACCCCTCACTGTCCTGCTGC 76  
73 CTGACTTTTCTCTGCAAGCTCCATGTTTTCATCTCTTCTTGGCTGTCTGCTGC 132  
77 TGTGCACTACTACTTGAAGTCACTTGAAGATGTTAAGTGAAGTGTGTA 136  
133 TGTGCTGCTACTACTTGAAGTCTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 192  
137 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196  
193 ATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252  
197 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
253 GGGGCAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312  
257 GAATATGCACTATCAAGATCCAGAGATCCAGGATTAAGGAGGATTAAGGAGGATTA 316  
313 GGAATGATTAATTTGAGATC---CAGATCTGCTTAATGAGGATTTCCGAAAGGAG 359  
317 ATGTGCTCTGATCAATTAAGATGATCTGTGATGATCAAGGAGGATCTGCTGAGGA 376  
370 ATGTGCTCTGATCAATTAAGATGATCTGTGATGATCAAGGAGGATCTGCTGAGGA 429



Oy	377	TACGTTCCCGGCTTATGAATGATATAAAATTTAGAACTGAAATTAGACATCAAGACG	436
Db	430	TCCAATCCCGAGGATATATGATATGATATAAAATTTAACTTGAACTTGTCATTAACCG	489
Oy	437	CCAAGTCACTCCAGCTCAGACTGECCTCATGCGGAGCTCTACATGAGCTTCTCCAAGACCC	496
Db	490	CCAAGTCACTCCCGAGCTCTGAGAGAGACTTCACTGACGCTTTCGAAGATCG	549
Oy	497	TAAACACGAGAGAAATG---TTGAGAGACACAGACACTGTGTACCTTCATATPAQC	553
Db	550	TTACCAACAGGGGACATGGCCACAGACACACGACACTGGGAGGCTCCCTGATATPA	609
Oy	554	ATGGAACAAAATTTCCAC-----ATGGCGTGAAGAA	586
Db	610	ATCTPAQCAAAATATCCAATTGGCCAAATGATTAAGGGACTCAATGGCCAAATGACT	669
Oy	587	TTAAGACTCTGAGAAACGATCGAAGCTGATCCACATTTGAGTGGAGTCTGTGCTG	646
Db	670	TACGGACTCTGAGGCAACCATGAAATAGGACATCTACATCGAGCAGGGATCTGTGCTG	729
Oy	647	GGTGACCTTGCGCACTTATCATTTGCTCTTAAATCCTTAAATGGATTTCTGTAGAAA	706
Db	730	GGCTGCTGTGGCTCTTATCTTGCGCGCTTAAATTTCAATGGTATTTCTATAGCAAG	789
Oy	707	AGAAGTTATGAGATTGAGCTTATTAACACTGGCCCACTTGCTCCAGAGGGTTGGCA	766
Db	790	AGAAAGTACAGAAATTTAAGCCTCATCTTTGGCCAACTCCCTCCTCAGGATTTGGCA	849
Oy	767	ATGCAGAGCAAGTCAAGATTGCGCTGTAGGAAATATCTACACATCGAGAGAAAGTAT	826
Db	850	ATGCAGTAGCAGAGGAAATTCGCTCAAGAGAAACATCTATACCATTTGAAGAAAGTAT	909
Oy	827	ATGAAGTGAAGAAATTCAAATGAGTACTACTACGTCAACAGCCAGACCACTCCGAC	886
Db	910	ATGAAGTGAAGAACCCCAATGAGTATTAATGCTATGTACAGAGAGGCAACCACTCCAC	969
Oy	887	CGCCTCTGAGCTGCCACTTTTAAAGC-----TCGCTTCATTTCTGACTTTGGT	936
Db	970	AACCTTTGGGGTGTGCGCTTTGGCAGMGGCANAARATCCAACCACTTAAATTTTGA	1029
Oy	937	ATTTCCTCTTTGTGAAAACATATGATATGCACTTGGSCAACCTCATTGGAGGTTGAC	996
Db	1030	TTTTTGTCTTTTTCGAAAMATATGAGCTGTGACCTG---ACTGGTTTGGRGKTYTGT	1085
Oy	997	CACAGCCACTGAGAAAGAGTTCCAGTTTCTGCGGAGTAATTTAA	1040
Db	1086	CMCTGCTATMTGRMGAGTTTTCCHTTTCAAAAAAAMATCMW	1129
RESULT 5			
BX526231			
ID	BX526231	standard; RNA; EST; 476 Bp.	
XX	BX526231;		
XX	BX526231.1		
SV	27-MAY-2003 (Rel. 75, Created)		
XX	27-MAY-2003 (Rel. 75, Last updated, Version 1)		
DT	RZPD Mus musculus cDNA clone IMAGP98J513740 = IMAGE:1478438 5' EST.		
XX	EST; expressed sequence tag.		
XX	Mus musculus (house mouse)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
XX	[1]		
RP	1-476		
RA	Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,		
RA	Korn B.;		
RT	;		

RJ	Submitted (28-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RJ	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
RJ	Feld 580, D-69120 Heidelberg, Germany
XX	
CC	RZPD; IMAGE998G53740.
CC	RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
CC	Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
CC	http://www.rzpd.de/Cloncards/cgi-bin/showlib.pl.cgi/response?libNo=981
CC	Contact: Ina Rolts
CC	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
CC	Heubenerweg 6, D-14059 Berlin, Germany
CC	Tel.: +49 30 32639 101
CC	Fax: +49 30 32639 111
CC	www.rzpd.de
CC	This clone is available royalty-free from RZPD;
CC	contact RZPD (clone@rzpd.de) for further information.
-CC	Seq primer: T7, Primer sequence: TAATACGACTCATTAAGGG
XX	
PH	Key
FT	Location/Qualifiers
FT	source
FT	1. .476
FT	/db_xref="taxon:10090"
FT	/note="1st strand cDNA was primed with a Not I - oligo(dT)
FT	primer [15' rgttaccatcatgaaatggagaggccgcctgaattctttttt
FT	TTTTTTT]TTTTT 3']; double-stranded cDNA was ligated to
FT	EcoRI adaptors 5'-AATTGGCAGAG-3' and 5'-CTGTGCCG-3'
FT	(pharmacia), digested with NotI and cloned into the NotI
FT	and EcoRI sites of the pTV73D-Pac vector. Library is
FT	normalized. Library was constructed by Beno Soares and M.
FT	Patiama Bernaldo."
FT	/organism="Mus musculus"
FT	/clone="IMAGE998G53740"
FT	/clone_1id="Soares NMPU pregnant uterus"
FT	/dev_stage="adult"
FT	/lab_host="DH10B"
XX	
SQ	Sequence 476 BP, 135 A; 124 C; 113 G; 104 T; 0 other;
	Query Match 16.5%; Score 446; DB 4; Length 476;
	Best Local Similarity 95.9%; Pred. No. 1.1e-58;
	Matches 468; Conservative 4; Mismatches 4; Indels 12; Gaps 1;
OY	1539 GAATCTATATTGTAATCGTAGACCCCACAATGAAGAAGCTAGGCTGTGTAAGACATGCTTGT 1598
Db	1 GAATCTATATTGTAATCGTAGACCCCACAATGAAAAAGCTAGGCTGTGTAAGACATGCTTGT 60
OY	1599 AGACTCAAGAAATGAGAGGTAAGAGCAACAACATCCCCGGGGCTTCGTCGACATGCAGC 165
Db	61 AGACTCAAGAAATGAGAGGTAAGAGCAACAACATCCCCGGGGCTTCGTCGACATGCAGC 120
OY	1659 TTAGCTAGAGTGCTGAGTTCCAATGCCAACAGATCCCTGCTCTAMAGTAAGATGGRCTG 1718
Db	121 TTAGCTAGAGTGCTGAGTTCCAATGCCAACAGATCCCTGCTCTAMAGTAAGATGGRCTG 180
OY	1719 AGTATCTGGCGCATGTCCTCAGTGGGGGTGTCTCTCTCTCAGANAAGACATGCATGWC 1778
Db	181 AGTATCTGGCGCATGTCCTCAGTGGGGGTGTCTCTCTCTCAGANAAGACATGCATGAAA 240
OY	i779 CCTGCATGAAATGA 1838
Db	241 CCTG-----CACACACACACACACACACACACACACACACACACATGAAATGA 288
OY	1839 AGGTTCTCTGTGCTGCTGCTACTCTCTATPAACATGTATCTCTACGAGACTCTCCTCTGC 1898
Db	289 AGGTTCTCTGTGCTGCTGCTACTCTCTATPAACATGTATCTCTACGAGACTCTCCTCTGC 348
OY	1899 CTCTGTTAAGACATGAGTGGAGAGATGGACAGCATGCCAGTAATTAATTCACACTCA 1958
Db	349 CTCTGTTAAGACATGAGTGGAGAGATGGACAGCATGCCAGTAATTAATTCACACTCA 408
OY	1959 GAAAGCTGGAGACAAAGCTGGAGAGATTCAAGAGACATGTGCCAACACATGCCAGACTCT 2018
Db	409 GAAAGCTGGAGACAAAGCTGGAGAGATTCAAGAGACATGTGCCAACACATGCCAGACTCT 468

QY 2019 TCTTACAC 2026  
 Db 469 TCTTACAC 476

RESULT 6  
 BB750705 440 bp mRNA linear EST 16-OCT-2001  
 LOCUS BB750705 RIKEN full-length enriched, pooled tissues, cerebellum,  
 etc. Mus musculus cDNA clone G130201N16 3', mRNA sequence.  
 ACCESSION BB750705 GI:16154941  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 440)

REFERENCE  
 AUTHORS Akimura,T., Araiawa,T., Carninci,P., Furuno,M., Hanagaki,T.,  
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,  
 Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,  
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
 Saito,R., Sakai,C., Sakai,K., Sakazume,N., Saeki,D., Sato,K.,  
 Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,  
 A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,  
 Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
 2001)

TITLE Unpublished  
 JOURNAL Contact: Yoshinide Hayashizaki  
 COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,  
 M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,  
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
 Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,  
 Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 Location/Qualifiers  
 1..440  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G130201N16"  
 /clone\_1lb="RIKEN full-length enriched, pooled tissues,  
 cerebellum, etc."  
 /note="pooled tissues; (tissue\_type=cerebellum,  
 dev\_stage=16 days neonate, sex=mixed),  
 (tissue\_type=cerebellum, dev\_stage=0 day neonate,  
 sex=mixed), (tissue\_type=hippocampus, dev\_stage=adult,  
 sex=male), (tissue\_type=whole body, dev\_stage=9 days

BASE COUNT 127 a 83 c 101 g 129 t  
 ORIGIN  
 Query Match 16.2%; Score 440; DB 10; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-58;  
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2228 GCAAAATATTCCTTGGAGAGACAGTCGTACCAAAATGATTTCCAGCCGGTGACCTCA 2287  
 1 GCAAAATATTCCTTGGAGAGACAGTCGTACCAAAATGATTTCCAGCCGGTGACCTCA 60

QY 2288 GTTTCATCTGGCTTACAGCTGCTGCCGCCCTTATCTGTGTGCTCCACTAT 2347  
 61 GTTTCATCTGGCTTACAGCTGCTGCCGCCCTTATCTGTGTGCTCCACTAT 120

QY 2348 AACAGAAATCAATTAATAGACCCGAGTGAATAATTAAGTAGACAGAAAGTACTTT 2407  
 121 AACAGAAATCAATTAATAGACCCGAGTGAATAATTAAGTAGACAGAAAGTACTTT 180

QY 2408 GTTCAAAAGATTTTTCATTTGGAGAGCAACTGTGTACATCAGAGACATCTTTAGTGA 246  
 181 GTTCAAAAGATTTTTCATTTGGAGAGCAACTGTGTACATCAGAGACATCTTTAGTGA 240

QY 2468 GGAACACCAAACTGTGGTACCGTTTTCATGATGAAATTTTGTGTTAGTTCCTTC 2527  
 241 GGAACACCAAACTGTGGTACCGTTTTCATGATGAAATTTTGTGTTAGTTCCTTC 300

QY 2528 TAGCTAGCTGTGGAGGCTGCTGCTTCTTAGGTGGTATGAAAGGAGACATCTTAACA 2587  
 301 TAGCTAGCTGTGGAGGCTGCTGCTTCTTAGGTGGTATGAAAGGAGACATCTTAACA 360

QY 2588 AATCCATTAGAGATTAACAGCTTCATGCAAGAGGAAATTAATCTCAATGTTTAAAG 2647  
 361 AATCCATTAGAGATTAACAGCTTCATGCAAGAGGAAATTAATCTCAATGTTTAAAG 420

QY 2648 TAATAAACTGTTACTGGCA 2667  
 421 TAATAAACTGTTACTGGCA 440

RESULT 7  
 BB7528967 381 bp mRNA linear EST 14-DEC-2002  
 LOCUS BB7528967 RIKEN full-length enriched, NOD-derived cDNA +ve  
 DEFINITION dendritic cells Mus musculus cDNA clone F630119G06 3', mRNA  
 sequence.  
 ACCESSION BB7528967  
 VERSION BB7528967.1 GI:26863346  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 381)

REFERENCE  
 AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 381)  
 Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,U., Bono,H., Kondo,S.,  
 Nikaide,I., Osato,N., Saito,R., Suzuki,H., Yamataka,I., Kiyosawa,H.,  
 Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,  
 Gojohori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,  
 Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,  
 Beisel,K.W., Blake,J.A., Bradt,D., Brueic,V., Chothia,C., Corbani,  
 L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.E., Forrest,  
 A., Frazer,K.S., Gastlerland,T., Gariboldi,M., Gissi,C., Godzik,A.,  
 Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,  
 Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M.,  
 King,B.L., Kongsawa,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,  
 P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,  
 H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,  
 Resole,G., Petrovsky,N., Piliat,R., Pontius,J.U., Qi,D.,  
 Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,  
 B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sempke,C.A., Setou,  
 M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale



BASE COUNT 106 a 97 c 86 g 85 t  
 ORIGIN

Query Match 12.8%; Score 345.6; DB 9; Length 374;  
 Best Local Similarity 95.1%; Pred. No. 2.8e-43;  
 Matches 367; Conservative 4; Mismatches 3; Indels 12; Gaps 1;

1540 AATCTATATTTGATTCGTAGACCCACATGAAAAAGCTGCTGTGAGACATGCTTGTGA 1599  
 1 AATCTATATTTGATTCGTAGACCCACATGAAAAAGCTGCTGTGAGACATGCTTGTGA 60  
 1600 GACTCAGAGATGAGAGATGAAAGGCAACAAGATCCCGGGGCTTGCTGACATGACGT 1659  
 61 GACTCAGAGATGAGAGATGAAAGGCAACAAGATCCCGGGGCTTGCTGACATGACGT 120  
 1660 TAGCTTAGGTGCTGAGTTCAGTCCAGTCCAGAGATCCCTGTCTCAGAGATGAGTGA 1719  
 121 TAGCTTAGGTGCTGAGTTCAGTCCAGTCCAGAGATCCCTGTCTCAGAGATGAGTGA 180  
 1720 GTATCTGGCGCATGTCCTATGCGGGGTTGTCTCTCTCTCTCAGAAAGACATGACATGCC 1779  
 181 GTATCTGGCGCATGTCCTATGCGGGGTTGTCTCTCTCTCTCAGAAAGACATGACATGAC 240  
 1780 CTGACATGAAATGAA 1839  
 241 CTG-----CACACACACACACACACACACACACACACACACACATGAAATGAA 288  
 1840 GATTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1899  
 289 GATTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348  
 1900 TCTGTTAAGCATGATGATGGAGCATG 1925  
 349 TCTGTTAAGCATGATGATGGAGCATG 374

RESULT 9 354 bp mRNA linear EST 30-SEP-1998  
 LOCUS A1152783  
 DEFINITION u91b01.x1 Soares\_NMPu Mus musculus cDNA clone IMAGE:1478185',  
 mRNA sequence.  
 ACCESSION A1152783  
 VERSION A1152783.1 GI:3681252  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 354)  
 Maria, M., Hillier, J., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Scheinberg, K., Stepec, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.

TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Maira M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu

This clone is available royalty-free through LINT; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:326561  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 330.  
 Location/Qualifiers

FEATURES  
 source 1..354  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

/db\_xref="taxon:10090"  
 /clone="IMAGE:1478185"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares\_NMPu"  
 /note="Organ: uterus; Vector: pUT73D-Pac (Pharmacia) with  
 a modified polylinker; 1st strand cDNA was prepared from  
 pregnant mouse uterus, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pUT73  
 vector. Library is normalized. Library was constructed by  
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 87 c 91 g 74 t  
 ORIGIN

Query Match 12.0%; Score 324; DB 9; Length 354;  
 Best Local Similarity 94.5%; Pred. No. 5.9e-40;  
 Matches 346; Conservative 4; Mismatches 4; Indels 12; Gaps 1;

1492 CTGAGATGAGGCTCGAGACATGAGTGCATTTGTGACAGACGAGAAATCTAATTTG 1551  
 1 CTGAGATGAGGCTCGAGACATGAGTGCATTTGTGACAGACGAGAAATCTAATTTG 60  
 1552 ATCTGAGACCCCATGAAAGCTAGGCTGCTGAGAGATGCTTGTAGATCTAAGAT 1611  
 61 ATCTGAGACCCCATGAAAGCTAGGCTGCTGAGAGATGCTTGTAGATCTAAGAT 120  
 1612 GGAGAGTAAAGGACACACAGATCCCGGGGCTTGCTGACATGACCTTAGCTAGTGC 1671  
 121 GGAGAGTAAAGGACACACAGATCCCGGGGCTTGCTGACATGACCTTAGCTAGTGC 180  
 1672 TGAGTTCCAAATGCCACAGAGTCCCTGCTCAGAAAGATGAGTATCTGCGCAA 1731  
 181 TGAGTTCCAAATGCCACAGAGTCCCTGCTCAGAAAGATGAGTATCTGCGCAA 240  
 1732 TGCTCAGTGGGGTGTCTCTCTCTCAGAAAGATGAGTATCTGCGCAA 1791  
 241 TGCTCAGTGGGGTGTCTCTCTCTCAGAAAGATGAGTATCTGCGCAA 291  
 1792 ACATGAAATGAGTCTCTGT 1851  
 292 ---CACACACACACACACACACACACACACACACACATGAAATGAGTCTCTGT 348  
 1852 GCCTGC 1857  
 349 GCCTGC 354

RESULT 10 583 bp mRNA linear EST 11-OCT-2002  
 LOCUS B0789118  
 DEFINITION i040b02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128595  
 3' similar to TR:054947 054947 KIDNEY INJURY MOLECULE-1 PRECURSOR  
 mRNA sequence.

ACCESSION B0789118  
 VERSION B0789118.1 GI:23838264  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 583)

Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,  
 Lemishka, I., Seacore, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Maria, M., Pape, D., Wylie, T., Martin, J., Blaisdin, A.,  
 Schmitt, A., Theising, B., Ritter, E., Renko, I., Bennett, J., Cardenas  
 M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R., Williams, T.,  
 Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished  
 TITLE JOURNAL  
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue



```

Db      735 AACCCAGCGCGCCAGGGAACCTGTCGCCGCTCTGGGGAAGAGGCTGTCCTG
QY      218 GGTCAACAGTGTACCAATGAGTTGCTGAACTGATGAAAGAAATGTCATATCAGAAAT
Db      675 TGTNTGAAATGGCAACGTGTGCTCAGGACTGATGAAAGGATGTAATTAATGGACAT
QY      278 CAGAGCATACCAAGGCGATGTCACAAAGAGATGTCCTGATCATTAAGA
Db      615 C---CAGATACGTGCTTAATGGGGATTTCCGCAAGAGATGTCCTTAACCATAGAGA
QY      338 ATGTGACTCTGATGACCATGAGACCTAATGCTGTCAGAGATACAGTTCCGTCTTAATGA
Db      558 ATGTGACTCTAGACAGACAGTGGGATCTACTGCTGTCGCCGATCCAAATCCAGGATATGA
QY      398 ATGATTAATAATTAAGAACTGAAATTTAGACATCAAGAGCCAAAGTCACTCCAGCTCAGA
Db      498 ATGATGAAATTAATTAACCTGAAAGTTGGTCAATCAACACAGACAG-----ACA
QY      458 CTGCCCATTGGGACTCTACTACAGCTTCTCCAGAACCCCTAACCCAGAGAGAAATGTT
Db      451 CAGACACTGGGGAGCTT-----CCGATATTAATCTAA
QY      518 CAGAGACACAGACACTGCTGACCTCCATATTAACATGGAACAAATTTCCACATGGG
Db      417 CACAAATATCCACATTGG-----CCATGAGTTACGGGACTCTAAGATTGG
QY      578 CTGATGAAATTAAGAGACTCTGAGAAACGATCAAGACTCTATCCATGAGTGGAG
Db      372 CCATGACTTACCGGACTCTGAGCAACCATCAAGAAATGAGCATCTACATGCGAGCGGA
QY      638 TCTCTGTGGGTTGACCCCTGCGCACTTATCATGTGTCCTTAATCTTAATGTAATCTCT
Db      312 TCTGTGTGGGCTGGCTGCTGCTCTTATCTTCCGCGCTTAATTTTAAATGATCTTC
QY      698 GTAGAGAAAGAAATTAATCAAGTTTGAAGCTTATTAACATGCGCAACTTCCCTCCAGAG
Db      252 ATAGCAAGAGAGATGATCAAAATTTAAGCTCATCTTTGGCAACCTCCCTCCAG
QY      758 GGTGGCAATTCAGAGACAGTCAAGATTCGCTGAGAAATATATCAACATGAGG
Db      192 GATTTGGCAATTCAGAGACAGTCAAGAGGAAATTCGCTCAGAAAGAAATATATCAACATGAGG
QY      818 AGAAGCATATGAGAGTGAAGTGAATTAATGATGATCTAATGCTAATGCTAATGCTAATGCTA
Db      132 AGAAGCATATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
QY      878 CATCTGACCGCTCTGAGACTGACACTTTAAAGG
Db      72 AACCTTCAACAACCTTTGGTGTGCTTGCATATGC

```

```

Tissue Procurement: Narayan Bhat
CDNA Library Preparation: CLOUTCH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: NDCM203 row: h column: 02
High quality sequence stop: 509.
Location/Qualifiers
1..863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30411529"
/issue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_11b="NIH MGC 191"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccatcaggcc);
Site 2: SfiI (ggcgcccgcc); Library is oligo-dT
primed and directionally cloned. PBMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulated
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CAAGGCGCATTTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTGAGGCGGAGCGGCCGACACATG-dT(30)BR-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
Kb (range 0.70-5.0 Kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."
BASE COUNT      221 a      222 c      203 g      215 t      2 others
ORIGIN
Query Match      10.8%; Score 291.8; DB 14; Length 863;
Best Local Similarity 72.4%; Pred. No. 3; 6e-35;
Matches 420; Conservative 0; Mismatches 153; Indels 7; Gaps 3;
QY      17 CGGATTTCCCTCCCAAGTATCATGTTTTCAGGTTTACCTCAACTGTCTGCTGC
Db      36 CTGACTTTTCTCTGCAAGCTCATGTTTTCACATCTTCTTGAAGTGTCTGCTGC
QY      77 TGCTGCACTACTACTGCAAGTCAATGAGTGAAGATGATTAAGTTGAGTTGATTA
Db      96 TGCTGCTACTACTACTTCAAGTCTTCAAGAGTCTTCAAGAGTGAATACAGGAGTGTCTCA
QY      137 ATGCTATCTGCTGCTGCAAGTCACTCTTCACTATCATCTGAGACACTTGTGCTATGCT
Db      156 ATGCTATCTGCTGCTGCTTCTTCAACCCAGCGCCGACGAGAACTCGTGGCTGTGCT
QY      197 GGGGCAAGGATTTCTGCTTGTGTCACAGTGAATGATGTTGCTCAGAACTGATGAA
Db      216 GGGGCAAGGAGCTGTCTGTTGATGAGTGAAGTGTGCTCAGAGCTGATGAA
QY      257 GAAATGTGACATATCAGAAATCCAGAGATACAGCTTAAAGGCGATCAACAAAGAG
Db      276 GGAATGTGATTTTGAATC---CAGATCTGCTTAATGAGATTTCCGAAAGAG
QY      317 ATGTGTCTGTATCATTAAGAAATGTGACTCTGAGTACCATGAGACTTGTCTGAGGA
Db      333 ATGTGTCTGTACATTAAGAAATGTGACTCTGAGACAGAGTGGAGTCTACTGTGCGGA
QY      377 TACAGTTCTGTGCTTATGATGATTAAGAAATTAAGTGAATTAAGATCAAGAG
Db      393 TCCAAATCCAGGCAATTAAGTGAAGAAATTAACCTGAAGTTGATCAACAG
QY      437 CCAAGGTCACTCAGCTCAGACTGCGGAGACTCTACTAGCTTCTCCAAAGACC
Db      453 CCAAGGTCACTCAGCTCAGACTCTGCAAGAGACTCTGCAAGCTTTCCAAAGATG
QY      497 TAAACACGAGAGAAATGTTGAG--AGACACAGACACT-GGTGACCTTCATTAATAC
Db      513 TTACACACAGGAGACATGCGCCAGACAGACACAGACACTGAGGAGGAGCTCCCTGATATA

```



QY 553 ATGGAACAAAATTTCCATGGCTGATGAATTAAG 592  
 DB 573 ATCTACACANATATCCATGGCCATGAAGTACGG 612

RESULT 13  
 LOCUS BQ581947  
 DEFINITION BQ581947 554 bp mRNA linear EST 20-JUN-2002  
 5' similar to TR:054947 054947 KIDNEY INJURY MOLECULE-1 PRECURSOR  
 ; mRNA sequence.

ACCESSION BQ581947  
 VERSION BQ581947  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 554)  
 Melton, D., Brown, J., Kently, G., Permut, A., Lee, C., Kaestner, K.,  
 Lemisha, I., Searce, M., Breslitz, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A.,  
 Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J., Cardenas,  
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,  
 Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished  
 Other ESTs: l11a10.x1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohpc.harvard.edu  
 Library was constructed by Dr. J. Ferrer in vivo mass-excised to  
 plasmid SK- by Dr. H. Inoue DNA sequencing by: Washington  
 University Genome Sequencing Center For information on obtaining a  
 clone please contact: Dr. Hiroshi Inoue (hinoue@wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 430.  
 Location/Qualifiers  
 1. 554  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6029490"  
 /issue\_type="Insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human Insulinoma"  
 /note="Organ: Pancreas; Vector: plasmid SK-; Site: 1;  
 XhoI; Site 2: EcoRI; Constructed with lambda ZapII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 plasmid SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/esc/lambda\_protocol.shtml).  
 Please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permut  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

BASE COUNT 141 a 147 c 138 g 128 t

ORIGIN

Query Match. 10.7%; Score 289.8; DB 13; Length 554;  
 Best Local Similarity 73.0%; Pred. No. 8.6e-35;  
 Matches 401; Conservative 0; Mismatches 142; Indels 6; Gaps 2;

QY 19 GATTTCCCTCCAGACTGATGTTTTCAGTCTTACCTCACTGTCGTCGTCG 78  
 DB 9 GACTTTCTTCTCAAGCTTCATGTTTTCACATCTTCCCTTGACTGTGTCGTCGTCG 68

QY 79 CTGCAACTACTACTTTCAGAGTCAATGGAAGTGTATTAAGTTGAGTTGTAATAAT 138  
 DB 69 CTGCTGCTACTACTTACAGAGTCTCAGAAAGTGAATACAGAGCGAGGTGCTGAGAA 128

QY 139 GCTATCTGCTCCGTCGAGTTACACTTACCTTACCTTGGGACACTTGTGCTTATGCTCG 198  
 DB 129 GCTATCTGCTCCGTCGTTCTACACCCAGCGCCCGGGAACCTGCTGCTGCTCG 188

QY 139 GCGAAGGAGTCTGCTGTCGTCAGAGTACCAATGAGTTCGAGACGATGAGATGAAGA 258  
 DB 189 GCGAAGGAGCTGCTGTCGTTGTAATGTCGCAAGTGTCTCAGACTATGAAGA 248

QY 259 AATGTGACATATCAGAAATCCAGAGATACAGTAAAGGCGATCTCAACAGAGAT 318  
 DB 249 GATGTGANTATTTGACATC---CAGATCTGGCTAAATGGGATTTCCGCAAGAGAT 305

QY 319 GTGCTCTGATCATTAAGATGTGACTGTGATGACCATGGACCTACTGTCGAGAT 378  
 DB 306 GTGCTCCGACCATAGAGATGTGACTGACGACAGTGGGATCTACTGCTGCGGATC 365

QY 379 CAGTCCCTGCTCTATGATGATTAATAATTAAGTGAATTTGACATCAAGACGCC 438  
 DB 366 CAAATCCAGCATATGATGATGATTAATAATTAAGTGAATTTGATGATCAACAGCC 425

QY 439 AAGTCACTCCAGCTCAGACTGCCCATGGGAGCTACTACTACGTTCTTCCAGAAACCTTA 498  
 DB 426 AAGTCACTCCGTCGACACGACTCTGACGAGAGACTTCACTGACGCTTTCCAGAGATCT 485

QY 499 ACCACGAGAGAAATG---TTACAGACACACGACACTGTGATCCCTCATTAACAT 555  
 DB 486 ACCACGAGGAGCATGCGCCAGCAGACACAGACACTGGGAGCTCCCTCATTAAT 545

QY 556 GGAACAAA 564  
 DB 546 CTAACACA 554

RESULT 14  
 LOCUS AU279844  
 DEFINITION AU279844 CHONS2 Homo sapiens cDNA clone CHONS2001704 5', mRNA  
 sequence.

ACCESSION AU279844  
 VERSION AU279844  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 575)  
 Imabayashi, H., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R., Isogai, T.,  
 Mori, T., Hata, J., Tomoya, Y., and Umezawa, A.  
 Redifferentiation of dedifferentiated chondrocytes and  
 chondrogenesis of human bone marrow stromal cells via chondrosphere  
 formation with an expression profiling by large-scale cDNA analysis  
 Unpublished  
 Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA Project, Sugiyama, T.; Wakamatsu, A.; Irie, R.; Umezawa,  
 A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.; Isomo,  
 Y.; Saito, K.; Nakamura, Y.; Masuo, Y.; Nagai, K.; Isogai, T.  
 HRI human cDNA project; cDNA library construction & 5'-end one  
 pass sequencing; Helix Research Institute.

FEATURES  
 source 1. 575  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"



/db\_xref="taxon:9606"  
 /clone="CHONS2001704"  
 /cell\_type="chondrocytes"  
 /clone\_lib="CHONS2"  
 /note="Vector: PME18SFL3"  
 BASE COUNT 146 a 149 c 144 g 136 t  
 ORIGIN

Query Match 10.6%; Score 286.8; DB 9; Length 575;  
 Best Local Similarity 72.9%; Pred. No. 2,4e-34;  
 Matches 388; Conservative 0; Mismatches 142; Indels 6; Gaps 2;

17 CGGATTTCCCTCCCAAGTACTGTTTTCAGTCTTACCTTCACTGTCCTGTCG 76  
 33 CTGACTTTCTTCTGCAAGCTCATGTTTTCATCTCTTCTGACTGTCTCTGTC 92  
 77 TGTGCAACTACTACTTTCAGAGTTCATGGAAGTGTATTAAGTTGAGTTGTA 136  
 93 TGTCTGTGCTACTACTTACAAAGTCTCAGAAAGTGAATACAGAGCGAGTGC 152  
 137 ATGCTATCTGCTCCCTGCAAGTTCATCTTACTCATCTGGGACACTTGTCT 196  
 153 ATGCTATCTGCTCCCTGCTTCTTACACCCGCGCCGAGGAACTCTGCTGCT 212  
 197 GGGGCAAGGAGTCTGCTTGTGTCACAGTGTACCAATGAGTGTCTCAGAACT 256  
 213 GGGGCAAGGAGTCTGCTTGTGTCACAGTGTACCAATGAGTGTCTCAGAACT 272  
 257 GAAATGTGACATTCAGAAATTCAGAGATACAGCTTAAAGGCGATCTCAAG 316  
 273 GGGATGTGATTTATGAGCATC---CAGATATGCTGTAATGAGGATTTCCG 359  
 317 ATGTGCTCTGATCATTAAGAAATGTGACTGTGATGACCTGGAACCTTACT 376  
 330 ATGTGCTCTGATCATTAAGAAATGTGACTGTGATGACCTGGAACCTTACT 389  
 377 TACAGTCTCTGCTTATGAATGATTAATAAATTAGAACTGAATTCATCAAG 436  
 390 TCCAAATCCAGCATTAATGATGATGAAAAATTTAACTGAAAGTGTCTATCA 449  
 437 CCAAGGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTG 496  
 450 CCAAGGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTG 509  
 497 TAAACCGAGAGAAATG---TTCAAGACACAGACACTGTGTGACCTTCAAT 553  
 510 TTACACACAGAGGAGCATGCGCCAGACAGACAGACAGACAGACAGACAG 569  
 554 ATGAA 559  
 570 ATCTAA 575

RESULT 15  
 CD522575 746 bp mRNA linear EST 06-JUN-2003  
 LOCUS AGENCOURT\_14353511 NIH\_MGC\_191 Homo sapiens CDNA clone  
 DEFINITION IMAGE:30411065 5', mRNA sequence.  
 ACCESSION CD522575  
 VERSION CD522575.1 GI:31454293  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
 1 (bases 1 to 746)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgaabs@mail.nih.gov  
 Tissue Procurement: Narayan Bhat  
 CDNA Library Preparation: CLOUTIER Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: NDCM202 row: d column: 18  
 High quality sequence stop: 592.  
 Location/Qualifiers

## FEATURES

source

1..746  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30411065"  
 /tissue\_type="Pooled"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 191"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc);  
 Site 2: SfiI (ggccgctccggc); Library is oligo-dT primed  
 and directionally cloned. PMBC - Peripheral Blood  
 Mononuclear Cells. RNA was pooled from 3/6hour stimulat  
 with PMA adn Ionomycin. 5' and 3' adaptors were used in  
 cloning as follows: 5' adaptor sequence:  
 5'-CACGGCATTAAGCC-3' and 3' adaptor sequence:  
 5'-ATTGAGGCGGAGCGGCGGCAATG-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.69  
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH MGC library."  
 BASE COUNT 191 a 185 c 192 g 177 t 1 others  
 ORIGIN

Query Match 10.2%; Score 276; DB 14; Length 746;  
 Best Local Similarity 73.1%; Pred. No. 1e-32;  
 Matches 383; Conservative 0; Mismatches 135; Indels 6; Gaps 2;

17 CGGATTTCCCTCCCAAGTACTGTTTTCAGTCTTACCTTCACTGTCCTGTCG 76  
 120 CTGACTTTCTTCTGCAAGCTCATGTTTTCATCTCTTCTGACTGTCTCTGTC 179  
 77 TGTGCAACTACTACTTTCAGAGTTCATGGAAGTGTATTAAGTTGAGTTGTA 136  
 180 TGTGCTGTACTACTTACAAAGTCTCAGAAAGTGAATACAGCGAGTGTGACA 239  
 137 ATGCTATCTGCTCCCTGCAAGTTCATCTTACTCAATCTGGGACACTTGTCT 196  
 240 ATGCTATCTGCTCCCTGCTTCTTACACCCGCGCCGAGGAACTCGTCTGCT 299  
 197 GGGGCAAGGAGTCTGCTTGTGTCACAGTGTACCAATGAGTGTCTCAGAACT 256  
 300 GGGGCAAGGAGTCTGCTTGTGTCACAGTGTACCAATGAGTGTCTCAGAACT 359  
 257 GAAATGTGACATTCAGAAATTCAGAGATACCAAGTGAAGGCGATCTCAAG 316  
 360 GGAATGTGATTAATGAGATC---CAGATCTGCTTAATGAGGATTTCCGAA 416  
 317 ATGTGCTCTGATCATTAAGAAATGAGTCTGAGATGACATGAGGACCTTCTG 376  
 417 ATGTGCTCTGACCATTAAGAAATGAGTCTGAGATGAGTGTGATCTCTGCG 476  
 477 TCCAAATCCAGCATTAATGAGTGAATTAATACTGAAGTGTGATCAACAG 536  
 437 CCAAGGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTG 496  
 537 CCAAGGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCTG 596  
 497 TAAACCGAGAGAAATG---TTCAAGACACAGACACTGTGCT 537

Mon Nov 24 13:30:30 2003

us-10-004-633-8.rst

Page 13

Db 597 TTACCACGAGGGGACATGGCCGACGACAGACACAGACTGGGG 640

Search completed: November 22, 2003, 01:46:24  
Job time : 5944.8 secs

**THIS PAGE BLANK (USPTO)**